

537, 121

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
17 June 2004 (17.06.2004)

PCT

(10) International Publication Number
WO 2004/050900 A2

- (51) International Patent Classification⁷: **C12Q**
- (21) International Application Number: PCT/US2003/040131
- (22) International Filing Date: 4 December 2003 (04.12.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/431,143 4 December 2002 (04.12.2002) US
60/431,206 4 December 2002 (04.12.2002) US
- (71) Applicant (*for all designated States except US*): **DI-ADEXUS, INC.** [US/US]; 343 Oyster Point Boulevard, South San Francisco, CA 94080 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **MACINA, Roberto, A.** [AR/US]; 4118 Crescendo Avenue, San Jose, CA 95136 (US). **TURNER, Leah, R.** [US/US]; 939 Rosette Court, Sunnyvale, CA 94086 (US). **SUN, Yongming** [US/US]; 551 Shoal Circle, Redwood City, CA 94065 (US). **RODRIGUEZ, Maria** [US/US]; 570 Avocet, #8109, Redwood City, CA 94065 (US).
- (74) Agents: **LICATA, Jane, Massey et al.**; Licata & Tyrrell P.C., 66 E. Main Street, Marlton, NJ 08053 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: COMPOSITIONS, SPLICE VARIANTS AND METHODS RELATING TO COLON SPECIFIC GENES AND PROTEINS

(57) Abstract: The present invention relates to newly identified nucleic acid molecules and polypeptides present in normal and neoplastic colon cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions containing the nucleic acid molecules, polypeptides, antibodies, agonists and antagonists of the invention and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating colon cancer and non-cancerous disease states in colon, identifying colon tissue, monitoring and identifying and/or designing agonists and antagonists of polypeptides of the invention. The uses also include gene therapy, production of transgenic animals and cells, and production of engineered colon tissue for treatment and research.



WO 2004/050900 A2

COMPOSITIONS, SPLICE VARIANTS AND METHODS RELATING TO COLON SPECIFIC GENES AND PROTEINS

5

INTRODUCTION

This application claims the benefit of priority from U.S. Provisional Patent Application Serial No. 60/431,143 filed December 4, 2002 and 60/431,206 filed December 4, 2002, which are herein incorporated by reference in their entireties.

FIELD OF THE INVENTION

10

The present invention relates to newly identified nucleic acids and polypeptides present in normal and neoplastic colon cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions comprising the nucleic acids, 15 polypeptides, antibodies, post translational modifications (PTMs), variants, derivatives, agonists and antagonists thereof and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating colon cancer and/or non-cancerous disease states in colon, identifying colon tissue and monitoring and identifying and/or designing agonists and antagonists of polypeptides of 20 the invention. The uses also include gene therapy, therapeutic molecules including but not limited to antibodies or antisense molecules, production of transgenic animals and cells, and production of engineered colon tissue for treatment and research.

BACKGROUND OF THE INVENTION

Colorectal cancer is the second most common cause of cancer death in the United 25 States and the third most prevalent cancer in both men and women. M. L. Davila & A. D. Davila, *Screening for Colon and Rectal Cancer*, in Colon and Rectal Cancer 47 (Peter S. Edelstein ed., 2000). The American Cancer Society estimates that there will be about 105,500 new cases of colon cancer and 42,000 new cases of rectal cancer in 2003 in the United States. Colon cancer and rectal cancer will cause about 57,100 deaths combined. 30 ACS Website: cancer.org on the world wide web. Nearly all cases of colorectal cancer arise from adenomatous polyps, some of which mature into large polyps, undergo abnormal growth and development, and ultimately progress into cancer. Davila at 55-56. This progression would appear to take at least 10 years in most patients, rendering it a

readily treatable form of cancer if diagnosed early, when the cancer is localized. Davila at 56; Walter J. Burdette, Cancer: Etiology, Diagnosis, and Treatment 125 (1998).

Although our understanding of the etiology of colon cancer is undergoing continual refinement, extensive research in this area points to a combination of factors, including age, hereditary and nonhereditary conditions, and environmental/dietary factors. Age is a key risk factor in the development of colorectal cancer, Davila at 48, with men and women over 40 years of age becoming increasingly susceptible to that cancer, Burdette at 126. Incidence rates increase considerably in each subsequent decade of life. Davila at 48. A number of hereditary and nonhereditary conditions have also been linked to a heightened risk of developing colorectal cancer, including familial adenomatous polyposis (FAP), hereditary nonpolyposis colorectal cancer (Lynch syndrome or HNPCC), a personal and/or family history of colorectal cancer or adenomatous polyps, inflammatory bowel disease, diabetes mellitus, and obesity. Davila at 47; Henry T. Lynch & Jane F. Lynch, *Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndromes)*, in Colon and Rectal Cancer 67-68 (Peter S. Edelstein ed., 2000).

Environmental/dietary factors associated with an increased risk of colorectal cancer include a high fat diet, intake of high dietary red meat, and sedentary lifestyle. Davila at 47; Reddy, B. S., *Prev. Med.* 16(4): 460-7 (1987). Conversely, environmental/dietary factors associated with a reduced risk of colorectal cancer include a diet high in fiber, folic acid, calcium, and hormone-replacement therapy in post-menopausal women. Davila at 50-55. The effect of antioxidants in reducing the risk of colon cancer is unclear. Davila at 53.

Because colon cancer is highly treatable when detected at an early, localized stage, screening should be a part of routine care for all adults starting at age 50, especially those with first-degree relatives with colorectal cancer. One major advantage of colorectal cancer screening over its counterparts in other types of cancer is its ability to not only detect precancerous lesions, but to remove them as well. Davila at 56. The key colorectal cancer screening tests in use today are fecal occult blood test, sigmoidoscopy, colonoscopy, double-contrast barium enema, and the carcinoembryonic antigen (CEA) test. Burdette at 125; Davila at 56.

The fecal occult blood test (FOBT) screens for colorectal cancer by detecting the amount of blood in the stool, the premise being that neoplastic tissue, particularly malignant tissue, bleeds more than typical mucosa, with the amount of bleeding increasing

with polyp size and cancer stage. Davila at 56-57. While effective at detecting early stage tumors, FOBT is unable to detect adenomatous polyps (pre-malignant lesions), and, depending on the contents of the fecal sample, is subject to rendering false positives. Davila at 56-59. Sigmoidoscopy and colonoscopy, by contrast, allow direct visualization of the bowel, and enable one to detect, biopsy, and remove adenomatous polyps. Davila at 59-60, 61. Despite the advantages of these procedures, there are accompanying downsides: sigmoidoscopy, by definition, is limited to the sigmoid colon and below, colonoscopy is a relatively expensive procedure, and both share the risk of possible bowel perforation and hemorrhaging. Davila at 59-60. Double-contrast barium enema (DCBE) enables detection of lesions better than FOBT, and almost as well as a colonoscopy, but it may be limited in evaluating the winding rectosigmoid region. Davila at 60. The CEA blood test, which involves screening the blood for carcinoembryonic antigen, shares the downside of FOBT, in that it is of limited utility in detecting colorectal cancer at an early stage. Burdette at 125.

Once colon cancer has been diagnosed, treatment decisions are typically made in reference to the stage of cancer progression. A number of techniques are employed to stage the cancer (some of which are also used to screen for colon cancer), including pathologic examination of resected colon, sigmoidoscopy, colonoscopy, and various imaging techniques. AJCC Cancer Staging Handbook 84 (Irvin D. Fleming et al. eds., 5th ed. 1998); Montgomery, R. C. and Ridge, J.A., *Semin. Surg. Oncol.* 15(3): 143-150 (1998). Moreover, chest films, liver functionality tests, and liver scans are employed to determine the extent of metastasis. Fleming at 84. While computerized tomography and magnetic resonance imaging are useful in staging colorectal cancer in its later stages, both have unacceptably low staging accuracy for identifying early stages of the disease, due to the difficulty that both methods have in (1) revealing the depth of bowel wall tumor infiltration and (2) diagnosing malignant adenopathy. Thoeni, R. F., *Radiol. Clin. N. Am.* 35(2): 457-85 (1997). Rather, techniques such as transrectal ultrasound (TRUS) are preferred in this context, although this technique is inaccurate with respect to detecting small lymph nodes that may contain metastases. David Blumberg & Frank G. Opelka, *Neoadjuvant and Adjuvant Therapy for Adenocarcinoma of the Rectum, in Colon and Rectal Cancer* 316 (Peter S. Edelstein ed., 2000).

Several classification systems have been devised to stage the extent of colorectal cancer, including the Dukes' system and the more detailed International Union against

Cancer-American Joint Committee on Cancer TNM staging system, which is considered by many in the field to be a more useful staging system. Burdette at 126-27. The TNM system, which is used for either clinical or pathological staging, is divided into four stages, each of which evaluates the extent of cancer growth with respect to primary tumor (T),
5 regional lymph nodes (N), and distant metastasis (M). Fleming at 84-85. The system focuses on the extent of tumor invasion into the intestinal wall, invasion of adjacent structures, the number of regional lymph nodes that have been affected, and whether distant metastasis has occurred. Fleming at 81.

Stage 0 is characterized by *in situ* carcinoma (Tis), in which the cancer cells are
10 located inside the glandular basement membrane (intraepithelial) or lamina propria (intramucosal). In this stage, the cancer has not spread to the regional lymph nodes (N0), and there is no distant metastasis (M0). In stage I, there is still no spread of the cancer to the regional lymph nodes and no distant metastasis, but the tumor has invaded the submucosa (T1) or has progressed further to invade the muscularis propria (T2). Stage II
15 also involves no spread of the cancer to the regional lymph nodes and no distant metastasis, but the tumor has invaded the subserosa, or the nonperitonealized pericolic or perirectal tissues (T3), or has progressed to invade other organs or structures, and/or has perforated the visceral peritoneum (T4). Stage III is characterized by any of the T substages, no distant metastasis, and either metastasis in 1 to 3 regional lymph nodes (N1)
20 or metastasis in four or more regional lymph nodes (N2). Lastly, stage IV involves any of the T or N substages, as well as distant metastasis. Fleming at 84-85; Burdette at 127.

Currently, pathological staging of colon cancer is preferable over clinical staging as pathological staging provides a more accurate prognosis. Pathological staging typically involves examination of the resected colon section, along with surgical examination of the
25 abdominal cavity. Fleming at 84. Clinical staging would be a preferred method of staging were it at least as accurate as pathological staging, as it does not depend on the invasive procedures of its counterpart.

Turning to the treatment of colorectal cancer, surgical resection results in a cure for roughly 50% of patients. Irradiation is used both preoperatively and postoperatively in
30 treating colorectal cancer. Chemotherapeutic agents, particularly 5-fluorouracil, are also powerful weapons in treating colorectal cancer. Other agents include irinotecan and floxuridine, cisplatin, levamisole, methotrexate, interferon- α , and leucovorin. Burdette at 125, 132-33. Nonetheless, thirty to forty percent of patients will develop a recurrence of

colon cancer following surgical resection, which in many patients is the ultimate cause of death. Wayne De Vos, *Follow-up After Treatment of Colon Cancer, Colon and Rectal Cancer* 225 (Peter S. Edelstein ed., 2000). Accordingly, colon cancer patients must be closely monitored to determine response to therapy and to detect persistent or recurrent disease and metastasis.

The next few paragraphs describe the some of molecular bases of colon cancer. In the case of FAP, the tumor suppressor gene APC (adenomatous polyposis coli), chromosomally located at 5q21, has been either inactivated or deleted by mutation. Alberts et al., *Molecular Biology of the Cell* 1288 (3d ed. 1994). The APC protein plays a role in a number of functions, including cell adhesion, apoptosis, and repression of the *c-myc* oncogene. N. R. Hall & R. D. Madoff, *Genetics and the Polyp-Cancer Sequence, Colon and Rectal Cancer* 8 (Peter S. Edelstein, ed., 2000). Of those patients with colorectal cancer who have normal APC genes, over 65% have such mutations in the cancer cells but not in other tissues. Alberts et al., *supra* at 1288. In the case of HPNCC, patients manifest abnormalities in the tumor suppressor gene HNPCC, but only about 15% of tumors contain the mutated gene. *Id.* A host of other genes have also been implicated in colorectal cancer, including the *K-ras*, *N-ras*, *H-ras* and *c-myc* oncogenes, and the tumor suppressor genes *DCC* (deleted in colon carcinoma) and *p53*. Hall & Madoff, at 8-9; Alberts et al., at 1288.

Abnormalities in Wg/Wnt signal transduction pathway are also associated with the development of colorectal carcinoma. Taipale, J. and Beachy, P.A. *Nature* 411: 349-354 (2001). Wnt1 is a secreted protein gene originally identified within mouse mammary cancers by its insertion into the mouse mammary tumor virus (MMTV) gene. The protein is homologous to the wingless (Wg) gene product of *Drosophila*, in which it functions as an important factor for the determination of dorsal-ventral segmentation and regulates the formation of fly imaginal discs. Wg/Wnt pathway controls cell proliferation, death and differentiation. Taipal (2001). There are at least 13 members in the Wnt family. These proteins have been found expressed mainly in the central nervous system (CNS) of vertebrates as well as other tissues such as mammary and intestine. The Wnt proteins are the ligands for a family of seven transmembrane domain receptors related to the Frizzled gene product in *Drosophila*. Binding Wnt to Frizzled stimulates the activity of the downstream target, Dishevelled, which in turn inactivates the glycogen synthetase kinase 3 β (GSK3 β). Taipal (2001). Usually active GSK3 β will form a complex with the

adenomatous polyposis coli (APC) protein and phosphorylate another complex member, β -catenin. Once phosphorylated, β -catenin is directed to degradation through the ubiquitin pathway. When GSK3 β or APC activity is down regulated, β -catenin is accumulated in the cytoplasm and binds to the T-cell factor or lymphocyte excitation factor (Tcf/Lef) family of transcriptional factors. Binding of β -catenin to Tcf releases the transcriptional repression and induces gene transcription. Among the genes regulated by β -catenin are a transcriptional repressor Engrailed, a transforming growth factor- β (TGF- β) family member Decapentaplegic, and the cytokine Hedgehog in *Drosophila*. β -Catenin is also involved in regulating cell adhesion by binding to α -catenin and E-cadherin. On the other hand, binding of β -catenin to these proteins controls the cytoplasmic β -catenin level and its complexing with TCF. Taipal (2001). Growth factor stimulation and activation of c-src or v-src also regulate β -catenin level by phosphorylation of α -catenin and its related protein, p120^{cas}. When phosphorylated, these proteins decrease their binding to E-cadherin and β -catenin resulting in the accumulation of cytoplasmic β -catenin. Reynolds, A.B. et al. *Mol. Cell Biol.* 14: 8333-8342 (1994). In colon cancer, c-src enzymatic activity has been shown to be increased to the level of v-src. Alternation of components in the Wg/Wnt pathway promotes colorectal carcinoma development. The best known modifications are to the APC gene. Nicola S et al. *Hum. Mol. Genet* 10:721-733 (2001). This germline mutation causes the appearance of hundreds to thousands of adenomatous polyps in the large bowel. It is the gene defect that accounts for the autosomally dominantly inherited FAP and related syndromes. The molecular alternations that occur in this pathway largely involve deletions of alleles of tumor-suppressor genes, such as APC, p53 and Deleted in Colorectal Cancer (DCC), combined with mutational activation of proto-oncogenes, especially c-Ki-ras. Aoki, T. et al. *Human Mutat.* 3: 342-346 (1994). All of these lead to genomic instability in colorectal cancers.

Another source of genomic instability in colorectal cancer is the defect of DNA mismatch repair (MMR) genes. Human homologues of the bacterial *mutHLS* complex (hMSH2, hMLH1, hPMS1, hPMS2 and hMSH6), which is involved in the DNA mismatch repair in bacteria, have been shown to cause the HNPCC (about 70-90% HNPCC) when mutated. Modrich, P. and Lahue, R. *Ann Rev. Biochem.* 65: 101-133 (1996); and Peltomäki, P. *Hum. Mol. Genet* 10: 735-740 (2001). The inactivation of these proteins leads to the accumulation of mutations and causes genetic instability that represents errors

in the accurate replication of the repetitive mono-, di-, tri- and tetra-nucleotide repeats, which are scattered throughout the genome (microsatellite regions). Jass, J.R. et al. *J. Gastroenterol Hepatol* 17: 17-26 (2002). Like in the classic FAP, mutational activation of c-Ki-ras is also required for the promotion of MSI in the alternative HNPCC. Mutations
5 in other proteins such as the tumor suppressor protein phosphatase PTEN (Zhou, X.P. et al. *Hum. Mol. Genet* 11: 445-450 (2002)), BAX (Buttler, L.M. *Aus. N. Z. J. Surg.* 69: 88-94 (1999)), Caspase-5 (Planck, M. *Cancer Genet Cytogenet.* 134: 46-54 (2002)), TGF β -R II (Fallik, D. et al. *Gastroenterol Clin Biol.* 24: 917-22 (2000)) and IGFII-R (Giovannucci E. *J. Nutr.* 131: 3109S-20S (2001)) have also been found in some colorectal
10 tumors possibly as the cause of MMR defect.

Some tyrosine kinases have been shown up-regulated in colorectal tumor tissues or cell lines like HT29. Skoudy, A. et al. *Biochem J.* 317 (Pt 1): 279-84 (1996). Focal adhesion kinase (FAK) and its up-stream kinase c-src and c-yes in colonic epithelial cells may play an important role in the promotion of colorectal cancers through the extracellular
15 matrix (ECM) and integrin-mediated signaling pathways. Jessup, J.M. et al., *The molecular biology of colorectal carcinoma, in: The Molecular Basis of Human Cancer*, 251-268 (Coleman W.B. and Tsongalis G.J. Eds. 2002). The formation of c-src/FAK complexes may coordinately deregulate VEGF expression and apoptosis inhibition. Recent evidences suggest that a specific signal-transduction pathway for cell survival that
20 implicates integrin engagement leads to FAK activation and thus activates PI-3 kinase and akt. In turn, akt phosphorylates BAD and blocks apoptosis in epithelial cells. The activation of c-src in colon cancer may induce VEGF expression through the hypoxia pathway. Other genes that may be implicated in colorectal cancer include Cox enzymes (Ota, S. et al. *Aliment Pharmacol. Ther.* 16 (Suppl 2): 102-106 (2002)), estrogen (al-
25 Azzawi, F. and Wahab, M. *Climacteric* 5: 3-14 (2002)), peroxisome proliferator-activated receptor- γ (PPAR- γ) (Gelman, L. et al. *Cell Mol. Life Sci.* 55: 932-943 (1999)), IGF-I (Giovannucci (2001)), thymine DNA glycosylase (TDG) (Hardeland, U. et al. *Prog. Nucleic Acid Res. Mol. Biol.* 68: 235-253 (2001)) and EGF (Mendelsohn, J. *Endocrine-Related Cancer* 8: 3-9 (2001)).

30 Gene deletion and mutation are not the only causes for development of colorectal cancers. Epigenetic silencing by DNA methylation also accounts for the loss of function of colorectal cancer suppressor genes. A strong association between MSI and CpG island methylation has been well characterized in sporadic colorectal cancers with high MSI but

not in those of hereditary origin. In one experiment, DNA methylation of MLH1, CDKN2A, MGMT, THBS1, RARB, APC, and p14ARF genes has been shown in 80%, 55%, 23%, 23%, 58%, 35%, and 50% of 40 sporadic colorectal cancers with high MSI respectively. Yamamoto, H. et al. *Genes Chromosomes Cancer* 33: 322-325 (2002); and
5 Kim, K.M. et al. *Oncogene*. 12;21(35): 5441-9 (2002). Carcinogen metabolism enzymes such as GST, NAT, CYP and MTHFR are also associated with an increased or decreased colorectal cancer risk. Pistorius, S. et al. *Kongressbd Dtsch Ges Chir Kongr* 118: 820-824 (2001); and Potter, J.D. *J. Natl. Cancer Inst.* 91: 916-932 (1999).

From the foregoing, it is clear that procedures used for detecting, diagnosing,
10 monitoring, staging, prognosticating, and preventing the recurrence of colorectal cancer are of critical importance to the outcome of the patient. Moreover, current procedures, while helpful in each of these analyses, are limited by their specificity, sensitivity, invasiveness, and/or their cost. As such, highly specific and sensitive procedures that would operate by way of detecting novel markers in cells, tissues, or bodily fluids, with
15 minimal invasiveness and at a reasonable cost, would be highly desirable.

Accordingly, there is a great need for more sensitive and accurate methods for predicting whether a person is likely to develop colorectal cancer, for diagnosing colorectal cancer, for monitoring the progression of the disease, for staging the colorectal cancer, for determining whether the colorectal cancer has metastasized, and for imaging
20 the colorectal cancer. Following accurate diagnosis, there is also a need for less invasive and more effective treatment of colorectal cancer.

Growth and metastasis of solid tumors are also dependent on angiogenesis. Folkman, J., 1986, *Cancer Research*, 46, 467-473; Folkman, J., 1989, *Journal of the National Cancer Institute*, 82, 4-6. It has been shown, for example, that tumors which
25 enlarge to greater than 2 mm must obtain their own blood supply and do so by inducing the growth of new capillary blood vessels. Once these new blood vessels become embedded in the tumor, they provide a means for tumor cells to enter the circulation and metastasize to distant sites such as liver, lung or bone. Weidner, N., et al., 1991, *The New England Journal of Medicine*, 324(1), 1-8.

30 Angiogenesis, defined as the growth or sprouting of new blood vessels from existing vessels, is a complex process that primarily occurs during embryonic development. The process is distinct from vasculogenesis, in that the new endothelial cells lining the vessel arise from proliferation of existing cells, rather than differentiating from

stem cells. The process is invasive and dependent upon proteolysis of the extracellular matrix (ECM), migration of new endothelial cells, and synthesis of new matrix components. Angiogenesis occurs during embryonic development of the circulatory system; however, in adult humans, angiogenesis only occurs as a response to a
5 pathological condition (except during the reproductive cycle in women).

Under normal physiological conditions in adults, angiogenesis takes place only in very restricted situations such as hair growth and wounding healing. Auerbach, W. and Auerbach, R., 1994, *Pharmacol Ther.* 63(3):265-311; Ribatti et al., 1991, *Haematologica* 76(4):311-20; Risau, 1997, *Nature* 386(6626):671-4. Angiogenesis progresses by a
10 stimulus which results in the formation of a migrating column of endothelial cells. Proteolytic activity is focused at the advancing tip of this "vascular sprout", which breaks down the ECM sufficiently to permit the column of cells to infiltrate and migrate. Behind the advancing front, the endothelial cells differentiate and begin to adhere to each other, thus forming a new basement membrane. The cells then cease proliferation and finally
15 define a lumen for the new arteriole or capillary.

Unregulated angiogenesis has gradually been recognized to be responsible for a wide range of disorders, including, but not limited to, cancer, cardiovascular disease, rheumatoid arthritis, psoriasis and diabetic retinopathy. Folkman, 1995, *Nat Med* 1(1):27-31; Isner, 1999, *Circulation* 99(13):1653-5; Koch, 1998, *Arthritis Rheum* 41(6):951-62;
20 Walsh, 1999, *Rheumatology* (Oxford) 38(2):103-12; Ware and Simons, 1997, *Nat Med* 3(2):158-64.

Of particular interest is the observation that angiogenesis is required by solid tumors for their growth and metastases. Folkman, 1986 *supra*; Folkman 1990, *J Natl. Cancer Inst.*, 82(1) 4-6; Folkman, 1992, *Semin Cancer Biol* 3(2):65-71; Zetter, 1998, *Annu*
25 *Rev Med* 49:407-24. A tumor usually begins as a single aberrant cell which can proliferate only to a size of a few cubic millimeters due to the distance from available capillary beds, and it can stay 'dormant' without further growth and dissemination for a long period of time. Some tumor cells then switch to the angiogenic phenotype to activate endothelial cells, which proliferate and mature into new capillary blood vessels. These newly formed
30 blood vessels not only allow for continued growth of the primary tumor, but also for the dissemination and recolonization of metastatic tumor cells. The precise mechanisms that control the angiogenic switch is not well understood, but it is believed that

neovascularization of tumor mass results from the net balance of a multitude of angiogenesis stimulators and inhibitors Folkman, 1995, *supra*.

One of the most potent angiogenesis inhibitors is endostatin identified by O'Reilly and Folkman. O'Reilly et al., 1997, *Cell* 88(2):277-85; O'Reilly et al., 1994, *Cell* 79(2):3
5 15-28. Its discovery was based on the phenomenon that certain primary tumors can inhibit the growth of distant metastases. O'Reilly and Folkman hypothesized that a primary tumor initiates angiogenesis by generating angiogenic stimulators in excess of inhibitors. However, angiogenic inhibitors, by virtue of their longer half life in the circulation, reach the site of a secondary tumor in excess of the stimulators. The net result is the growth of
10 primary tumor and inhibition of secondary tumor. Endostatin is one of a growing list of such angiogenesis inhibitors produced by primary tumors. It is a proteolytic fragment of a larger protein: endostatin is a 20 kDa fragment of collagen XVIII (amino acid H1132-K1315 in murine collagen XVIII). Endostatin has been shown to specifically inhibit endothelial cell proliferation in vitro and block angiogenesis in vivo. More importantly,
15 administration of endostatin to tumor-bearing mice leads to significant tumor regression, and no toxicity or drug resistance has been observed even after multiple treatment cycles. Boehm et al., 1997, *Nature* 390(6658):404-407. The fact that endostatin targets genetically stable endothelial cells and inhibits a variety of solid tumors makes it a very attractive candidate for anticancer therapy. Fidler and Ellis, 1994, *Cell* 79(2):185-8; Gastl et al.,
20 1997, *Oncology* 54(3):177-84; Hinsbergh et al., 1999, *Ann Oncol* 10 Suppl 4:60-3. In addition, angiogenesis inhibitors have been shown to be more effective when combined with radiation and chemotherapeutic agents. Klement, 2000, *J. Clin Invest*, 105(8) R15-24. Browder, 2000, *Cancer Res.* 6-(7) 1878-86, Arap et al., 1998, *Science* 279(5349):377-80; Mauceri et al., 1998, *Nature* 394(6690):287-91.

25

SUMMARY OF THE INVENTION

The present invention solves many needs in the art by providing nucleic acid molecules, polypeptides and antibodies thereto, variants and derivatives of the nucleic acids and polypeptides, and agonists and antagonists thereto that may be used to identify, diagnose, monitor, stage, image and treat colon cancer and/or non-cancerous disease states
30 in colon; identify and monitor colon tissue; and identify and design agonists and antagonists of polypeptides of the invention. The invention also provides gene therapy,

methods for producing transgenic animals and cells, and methods for producing engineered colon tissue for treatment and research.

One aspect of the present invention relates to nucleic acid molecules that are specific to colon cells, colon tissue and/or the colon organ. These colon specific nucleic acids (CSNAs) may be a naturally occurring cDNA, genomic DNA, RNA, or a fragment of one of these nucleic acids, or may be a non-naturally occurring nucleic acid molecule. If the CSNA is genomic DNA, then the CSNA is a colon specific gene (CSG). If the CSNA is RNA, then it is a colon specific transcript encoded by a CSG. Due to alternative splicing and transcriptional modification one CSG may encode for multiple colon specific RNAs. In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that is specific to colon. More preferred is a nucleic acid molecule that encodes a polypeptide comprising an amino acid sequence of SEQ ID NO: 113-259. In another preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-112. For the CSNA sequences listed herein, DEX0449_001.nt.1 corresponds to SEQ ID NO: 1. For sequences with multiple splice variants, the parent sequence DEX0449_001.nt.1, will be followed by DEX0449_001.nt.2, etc. for each splice variant. The sequences of the corresponding peptides are listed as DEX0449_001.aa.1, etc. For the mapping of all of the nucleotides and peptides, see the table in the Example 1 section below.

This aspect of the present invention also relates to nucleic acid molecules that selectively hybridize or exhibit substantial sequence similarity to nucleic acid molecules encoding a Colon Specific Protein (CSP), or that selectively hybridize or exhibit substantial sequence similarity to a CSNA. In one embodiment of the present invention the nucleic acid molecule comprises an allelic variant of a nucleic acid molecule encoding a CSP, or an allelic variant of a CSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid sequence that encodes a CSP or a part of a nucleic acid sequence of a CSNA.

In addition, this aspect of the present invention relates to a nucleic acid molecule further comprising one or more expression control sequences controlling the transcription and/or translation of all or a part of a CSNA or the transcription and/or translation of a nucleic acid molecule that encodes all or a fragment of a CSP.

Another aspect of the present invention relates to vectors and/or host cells comprising a nucleic acid molecule of this invention. In a preferred embodiment, the

nucleic acid molecule of the vector and/or host cell encodes all or a fragment of a CSP. In another preferred embodiment, the nucleic acid molecule of the vector and/or host cell comprises all or a part of a CSNA. Vectors and host cells of the present invention are useful in the recombinant production of polypeptides, particularly CSPs of the present invention.

Another aspect of the present invention relates to polypeptides encoded by a nucleic acid molecule of this invention. The polypeptide may comprise either a fragment or a full-length protein. In a preferred embodiment, the polypeptide is a CSP. However, this aspect of the present invention also relates to mutant proteins (muteins) of CSPs, fusion proteins of which a portion is a CSP, and proteins and polypeptides encoded by allelic variants of a CSNA as provided herein.

A further aspect of the present invention is a novel splice variant which encodes an amino acid sequence that provides a novel region to be targeted for the generation of reagents that can be used in the detection and/or treatment of cancer. The novel amino acid sequence may lead to a unique protein structure, protein subcellular localization, biochemical processing or function. This information can be used to directly or indirectly facilitate the generation of additional or novel therapeutics or diagnostics. The nucleotide sequence in this novel splice variant can be used as a nucleic acid probe for the diagnosis and/or treatment of cancer.

Another aspect of the present invention relates to antibodies and other binders that specifically bind to a polypeptide of the instant invention. Accordingly antibodies or binders of the present invention specifically bind to CSPs, muteins, fusion proteins, and/or homologous proteins or polypeptides encoded by allelic variants of a CSNA as provided herein.

Another aspect of the present invention relates to agonists and antagonists of the nucleic acid molecules and polypeptides of this invention. The agonists and antagonists of the instant invention may be used to treat colon cancer and non-cancerous disease states in colon and to produce engineered colon tissue.

Another aspect of the present invention relates to methods for using the nucleic acid molecules to detect or amplify nucleic acid molecules that have similar or identical nucleic acid sequences compared to the nucleic acid molecules described herein. Such methods are useful in identifying, diagnosing, monitoring, staging, imaging and treating colon cancer and/or non-cancerous disease states in colon. Such methods are also useful

in identifying and/or monitoring colon tissue. In addition, measurement of levels of one or more of the nucleic acid molecules of this invention may be useful as a diagnostic as part of a panel in combination with known other markers, particularly those described in the colon cancer background section above.

5 Another aspect of the present invention relates to use of the nucleic acid molecules of this invention in gene therapy, for producing transgenic animals and cells, and for producing engineered colon tissue for treatment and research.

 Another aspect of the present invention relates to methods for detecting polypeptides of this invention, preferably using antibodies thereto. Such methods are
10 useful to identify, diagnose, monitor, stage, image and treat colon cancer and non-cancerous disease states in colon. In addition, measurement of levels of one or more of the polypeptides of this invention may be useful to identify, diagnose, monitor, stage, and/or image colon cancer in combination with known other markers, particularly those described in the colon cancer background section above. The polypeptides of the present
15 invention can also be used to identify and/or monitor colon tissue, and to produce engineered colon tissue.

 Yet another aspect of the present invention relates to a computer readable means of storing the nucleic acid and amino acid sequences of the invention. The records of the computer readable means can be accessed for reading and displaying of sequences for
20 comparison, alignment and ordering of the sequences of the invention to other sequences. In addition, the computer records regarding the nucleic acid and/or amino acid sequences and/or measurements of their levels may be used alone or in combination with other markers to diagnose colon related diseases.

DETAILED DESCRIPTION OF THE INVENTION

25 Definitions and General Techniques

 Unless otherwise defined herein, scientific and technical terms used in connection with the present invention shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. Generally,
30 nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well known and commonly used in

the art. The methods and techniques of the present invention are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. *See, e.g.,* Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Press (2001); Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2000); Ausubel *et al.*, Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology – 4th Ed., Wiley & Sons (1999); Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1990); and Harlow and Lane, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1999).

Enzymatic reactions and purification techniques are performed according to manufacturer's specifications, as commonly accomplished in the art or as described herein. The nomenclatures used in connection with, and the laboratory procedures and techniques of, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well known and commonly used in the art. Standard techniques are used for chemical syntheses, chemical analyses, pharmaceutical preparation, formulation, and delivery, and treatment of patients.

The following terms, unless otherwise indicated, shall be understood to have the following meanings:

A "nucleic acid molecule" of this invention refers to a polymeric form of nucleotides and includes both sense and antisense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers of the above. A nucleotide refers to a ribonucleotide, deoxynucleotide or a modified form of either type of nucleotide. A "nucleic acid molecule" as used herein is synonymous with "nucleic acid" and "polynucleotide." The term "nucleic acid molecule" usually refers to a molecule of at least 10 bases in length, unless otherwise specified. The term includes single- and double-stranded forms of DNA. In addition, a polynucleotide may include either or both naturally occurring and modified nucleotides linked together by naturally occurring and/or non-naturally occurring nucleotide linkages.

Nucleotides are represented by single letter symbols in nucleic acid molecule sequences. The following table lists symbols identifying nucleotides or groups of

nucleotides which may occupy the symbol position on a nucleic acid molecule. See Nomenclature Committee of the International Union of Biochemistry (NC-IUB), Nomenclature for incompletely specified bases in nucleic acid sequences, Recommendations 1984., *Eur J Biochem.* 150(1):1-5 (1985).

Symbol	Meaning	Group/Origin of Designation	Complementary Symbol
a	a	Adenine	t/u
g	g	Guanine	c
c	c	Cytosine	g
t	t	Thymine	a
u	u	Uracil	a
r	g or a	puRine	Y
y	t/u or c	pYrimidine	r
m	a or c	aMino	k
k	g or t/u	Keto	m
s	g or c	Strong interactions 3H-bonds	w
w	a or t/u	Weak interactions 2H-bonds	s
b	g or c or t/u	not a	v
d	a or g or t/u	not c	h
h	a or c or t/u	not g	d
v	a or g or c	not t, not u	b
n	a or g or c or t/u, unknown, or other	aNy	n

5

The nucleic acid molecules may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.), pendent moieties (*e.g.*, polypeptides), intercalators (*e.g.*, acridine, psoralen, etc.), chelators, alkylators, and modified linkages (*e.g.*, alpha anomeric nucleic acids, etc.) The term "nucleic acid molecule" also includes any topological conformation, including single-stranded, double-stranded, partially duplexed, triplexed, hairpinned, circular and padlocked conformations. Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

20

A "gene" is defined as a nucleic acid molecule that comprises a nucleic acid sequence that encodes a polypeptide and the expression control sequences that surround the nucleic acid sequence that encodes the polypeptide. For instance, a gene may comprise a promoter, one or more enhancers, a nucleic acid sequence that encodes a polypeptide, downstream regulatory sequences and, possibly, other nucleic acid sequences involved in regulation of the expression of an RNA. As is well known in the art, eukaryotic genes usually contain both exons and introns. The term "exon" refers to a nucleic acid sequence found in genomic DNA that is bioinformatically predicted and/or experimentally confirmed to contribute contiguous sequence to a mature mRNA transcript. The term "intron" refers to a nucleic acid sequence found in genomic DNA that is predicted and/or confirmed to not contribute to a mature mRNA transcript, but rather to be "spliced out" during processing of the transcript.

A nucleic acid molecule or polypeptide is "derived" from a particular species if the nucleic acid molecule or polypeptide has been isolated from the particular species, or if the nucleic acid molecule or polypeptide is homologous to a nucleic acid molecule or polypeptide isolated from a particular species.

An "isolated" or "substantially pure" nucleic acid or polynucleotide (*e.g.*, an RNA, DNA or a mixed polymer) is one which is substantially separated from other cellular components that naturally accompany the native polynucleotide in its natural host cell, *e.g.*, ribosomes, polymerases, or genomic sequences with which it is naturally associated. The term embraces a nucleic acid or polynucleotide that (1) has been removed from its naturally occurring environment, (2) is not associated with all or a portion of a polynucleotide in which the "isolated polynucleotide" is found in nature, (3) is operatively linked to a polynucleotide which it is not linked to in nature, (4) does not occur in nature as part of a larger sequence or (5) includes nucleotides or internucleoside bonds that are not found in nature. The term "isolated" or "substantially pure" also can be used in reference to recombinant or cloned DNA isolates, chemically synthesized polynucleotide analogs, or polynucleotide analogs that are biologically synthesized by heterologous systems. The term "isolated nucleic acid molecule" includes nucleic acid molecules that are integrated into a host cell chromosome at a heterologous site, recombinant fusions of a native fragment to a heterologous sequence, recombinant vectors present as episomes or as integrated into a host cell chromosome.

A "part" of a nucleic acid molecule refers to a nucleic acid molecule that comprises a partial contiguous sequence of at least 10 bases of the reference nucleic acid molecule. Preferably, a part comprises at least 15 to 20 bases of a reference nucleic acid molecule. In theory, a nucleic acid sequence of 17 nucleotides is of sufficient length to occur at random less frequently than once in the three gigabase human genome, and thus provides a nucleic acid probe that can uniquely identify the reference sequence in a nucleic acid mixture of genomic complexity. A preferred part is one that comprises a nucleic acid sequence that can encode at least 6 contiguous amino acid sequences (fragments of at least 18 nucleotides) because they are useful in directing the expression or synthesis of peptides that are useful in mapping the epitopes of the polypeptide encoded by the reference nucleic acid. See, e.g., Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984); and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. A part may also comprise at least 25, 30, 35 or 40 nucleotides of a reference nucleic acid molecule, or at least 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides of a reference nucleic acid molecule. A part of a nucleic acid molecule may comprise no other nucleic acid sequences. Alternatively, a part of a nucleic acid may comprise other nucleic acid sequences from other nucleic acid molecules.

The term "oligonucleotide" refers to a nucleic acid molecule generally comprising a length of 200 bases or fewer. The term often refers to single-stranded deoxyribonucleotides, but it can refer as well to single-or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others. Preferably, oligonucleotides are 10 to 60 bases in length and most preferably 12, 13, 14, 15, 16, 17, 18, 19 or 20 bases in length. Other preferred oligonucleotides are 25, 30, 35, 40, 45, 50, 55 or 60 bases in length. Oligonucleotides may be single-stranded, e.g. for use as probes or primers, or may be double-stranded, e.g. for use in the construction of a mutant gene. Oligonucleotides of the invention can be either sense or antisense oligonucleotides. An oligonucleotide can be derivatized or modified as discussed above for nucleic acid molecules.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by

expression of DNAs in cells and organisms. Initially, chemically synthesized DNAs typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such
5 oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP. The 3' end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be
10 prevented selectively, where desired, by removing the 5' phosphates of the other polynucleotide(s) prior to ligation.

The term "naturally occurring nucleotide" referred to herein includes naturally occurring deoxyribonucleotides and ribonucleotides. The term "modified nucleotides" referred to herein includes nucleotides with modified or substituted sugar groups and the
15 like. The term "nucleotide linkages" referred to herein includes nucleotide linkages such as phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoraniladate, phosphoroamidate, and the like. *See e.g., LaPlanche et al. Nucl. Acids Res.* 14:9081-9093 (1986); *Stein et al. Nucl. Acids Res.* 16:3209-3221 (1988); *Zon et al. Anti-Cancer Drug Design* 6:539-568 (1991); *Zon et al.,*
20 *in Eckstein (ed.) Oligonucleotides and Analogues: A Practical Approach*, pp. 87-108, Oxford University Press (1991); *Uhlmann and Peyman Chemical Reviews* 90:543 (1990), and U.S. Patent No. 5,151,510, the disclosure of which is hereby incorporated by reference in its entirety.

Unless specified otherwise, the left hand end of a polynucleotide sequence in sense
25 orientation is the 5' end and the right hand end of the sequence is the 3' end. In addition, the left hand direction of a polynucleotide sequence in sense orientation is referred to as the 5' direction, while the right hand direction of the polynucleotide sequence is referred to as the 3' direction. Further, unless otherwise indicated, each nucleotide sequence is set forth herein as a sequence of deoxyribonucleotides. It is intended, however, that the given
30 sequence be interpreted as would be appropriate to the polynucleotide composition: for example, if the isolated nucleic acid is composed of RNA, the given sequence intends ribonucleotides, with uridine substituted for thymidine.

The term "allelic variant" refers to one of two or more alternative naturally occurring forms of a gene, wherein each gene possesses a unique nucleotide sequence. In a preferred embodiment, different alleles of a given gene have similar or identical biological properties.

5 The term "percent sequence identity" in the context of nucleic acid sequences refers to the residues in two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over a stretch of at least about nine nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about
10 32 nucleotides, and preferably at least about 36 or more nucleotides. There are a number of different algorithms known in the art which can be used to measure nucleotide sequence identity. For instance, polynucleotide sequences can be compared using FASTA, Gap or Bestfit, which are programs in Wisconsin Package Version 10.0, Genetics Computer
15 Group (GCG), Madison, Wisconsin. FASTA, which includes, *e.g.*, the programs FASTA2 and FASTA3, provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson, *Methods Enzymol.* 183: 63-98 (1990); Pearson, *Methods Mol. Biol.* 132: 185-219 (2000); Pearson, *Methods Enzymol.* 266: 227-258 (1996); Pearson, *J. Mol. Biol.* 276: 71-84 (1998)). Unless otherwise specified, default parameters for a particular program or algorithm are used. For instance,
20 percent sequence identity between nucleic acid sequences can be determined using FASTA with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) or using Gap with its default parameters as provided in GCG Version 6.1.

A reference to a nucleic acid sequence encompasses its complement unless otherwise specified. Thus, a reference to a nucleic acid molecule having a particular
25 sequence should be understood to encompass its complementary strand, with its complementary sequence. The complementary strand is also useful, *e.g.*, for antisense therapy, double-stranded RNA (dsRNA) inhibition (RNAi), combination of triplex and antisense, hybridization probes and PCR primers.

In the molecular biology art, researchers use the terms "percent sequence identity",
30 "percent sequence similarity" and "percent sequence homology" interchangeably. In this application, these terms shall have the same meaning with respect to nucleic acid sequences only.

The term “substantial similarity” or “substantial sequence similarity,” when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 50%, more preferably 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80%, preferably at least about 90%, and more preferably at least about 95-98% of the nucleotide bases, as measured by any well known algorithm of sequence identity, such as FASTA, BLAST or Gap, as discussed above.

Alternatively, substantial similarity exists between a first and second nucleic acid sequence when the first nucleic acid sequence or fragment thereof hybridizes to an antisense strand of the second nucleic acid, under selective hybridization conditions. Typically, selective hybridization will occur between the first nucleic acid sequence and an antisense strand of the second nucleic acid sequence when there is at least about 55% sequence identity between the first and second nucleic acid sequences— preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90% — over a stretch of at least about 14 nucleotides, more preferably at least 17 nucleotides, even more preferably at least 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 or 100 nucleotides.

Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, solvents, the base composition of the hybridizing species, length of the complementary regions, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. “Stringent hybridization conditions” and “stringent wash conditions” in the context of nucleic acid hybridization experiments depend upon a number of different physical parameters. The most important parameters include temperature of hybridization, base composition of the nucleic acids, salt concentration and length of the nucleic acid. One having ordinary skill in the art knows how to vary these parameters to achieve a particular stringency of hybridization. In general, “stringent hybridization” is performed at about 25°C below the thermal melting point (T_m) for the specific DNA hybrid under a particular set of conditions. “Stringent washing” is performed at temperatures about 5°C lower than the T_m for the specific DNA hybrid under a particular set of conditions. The T_m is the temperature at which 50% of the target sequence hybridizes to a perfectly matched probe. See Sambrook (1989), *supra*, p. 9.51.

The T_m for a particular DNA-DNA hybrid can be estimated by the formula:

$T_m = 81.5^{\circ}\text{C} + 16.6 (\log_{10}[\text{Na}^+]) + 0.41 (\text{fraction G} + \text{C}) - 0.63 (\% \text{ formamide}) - (600/l)$ where l is the length of the hybrid in base pairs.

The T_m for a particular RNA-RNA hybrid can be estimated by the formula:

$$T_m = 79.8^{\circ}\text{C} + 18.5 (\log_{10}[\text{Na}^+]) + 0.58 (\text{fraction G} + \text{C}) + 11.8 (\text{fraction G} + \text{C})^2 - 0.35 (\% \text{ formamide}) - (820/l).$$

The T_m for a particular RNA-DNA hybrid can be estimated by the formula:

$$T_m = 79.8^{\circ}\text{C} + 18.5 (\log_{10}[\text{Na}^+]) + 0.58 (\text{fraction G} + \text{C}) + 11.8 (\text{fraction G} + \text{C})^2 - 0.50 (\% \text{ formamide}) - (820/l).$$

In general, the T_m decreases by 1-1.5°C for each 1% of mismatch between two nucleic acid sequences. Thus, one having ordinary skill in the art can alter hybridization and/or washing conditions to obtain sequences that have higher or lower degrees of sequence identity to the target nucleic acid. For instance, to obtain hybridizing nucleic acids that contain up to 10% mismatch from the target nucleic acid sequence, 10-15°C would be subtracted from the calculated T_m of a perfectly matched hybrid, and then the hybridization and washing temperatures adjusted accordingly. Probe sequences may also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization conditions are well known in the art.

An example of stringent hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100 complementary residues on a filter in a Southern or Northern blot or for screening a library is 50% formamide/6X SSC at 42°C for at least ten hours and preferably overnight (approximately 16 hours). Another example of stringent hybridization conditions is 6X SSC at 68°C without formamide for at least ten hours and preferably overnight. An example of moderate stringency hybridization conditions is 6X SSC at 55°C without formamide for at least ten hours and preferably overnight. An example of low stringency hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100 complementary residues on a filter in a Southern or northern blot or for screening a library is 6X SSC at 42°C for at least ten hours. Hybridization conditions to identify nucleic acid sequences that are similar but not identical can be identified by experimentally changing the hybridization temperature from 68°C to 42°C while keeping the salt concentration constant (6X SSC), or keeping the hybridization temperature and salt concentration constant (*e.g.* 42°C and 6X SSC) and varying the formamide concentration from 50% to

0%. Hybridization buffers may also include blocking agents to lower background. These agents are well known in the art. *See* Sambrook *et al.* (1989), *supra*, pages 8.46 and 9.46-9.58. *See also* Ausubel (1992), *supra*, Ausubel (1999), *supra*, and Sambrook (2001), *supra*.

5 Wash conditions also can be altered to change stringency conditions. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see* Sambrook (1989), *supra*, for SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove excess probe. An exemplary medium stringency wash for duplex DNA of more than 100 base pairs is 1x SSC at 45°C for 15 minutes. An
10 exemplary low stringency wash for such a duplex is 4x SSC at 40°C for 15 minutes. In general, signal-to-noise ratio of 2x or higher than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

As defined herein, nucleic acids that do not hybridize to each other under stringent conditions are still substantially similar to one another if they encode polypeptides that are
15 substantially identical to each other. This occurs, for example, when a nucleic acid is created synthetically or recombinantly using a high codon degeneracy as permitted by the redundancy of the genetic code.

Hybridization conditions for nucleic acid molecules that are shorter than 100 nucleotides in length (*e.g.*, for oligonucleotide probes) may be calculated by the formula:
20 $T_m = 81.5^{\circ}\text{C} + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\text{fraction G+C}) - (600/\text{N})$, wherein N is change length and the $[\text{Na}^+]$ is 1 M or less. *See* Sambrook (1989), *supra*, p. 11.46. For hybridization of probes shorter than 100 nucleotides, hybridization is usually performed under stringent conditions (5-10°C below the T_m) using high concentrations (0.1-1.0 pmol/ml) of probe. *Id.* at p. 11.45. Determination of hybridization using mismatched
25 probes, pools of degenerate probes or “guessmers,” as well as hybridization solutions and methods for empirically determining hybridization conditions are well known in the art. *See, e.g.*, Ausubel (1999), *supra*; Sambrook (1989), *supra*, pp. 11.45-11.57.

The term “digestion” or “digestion of DNA” refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The
30 various restriction enzymes referred to herein are commercially available and their reaction conditions, cofactors and other requirements for use are known and routine to the skilled artisan. For analytical purposes, typically, 1 µg of plasmid or DNA fragment is digested with about 2 units of enzyme in about 20 µl of reaction buffer. For the purpose of

isolating DNA fragments for plasmid construction, typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes. Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and are specified by commercial suppliers. Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may be purified by electrophoresis through an agarose or polyacrylamide gel, using well known methods that are routine for those skilled in the art.

10 The term "ligation" refers to the process of forming phosphodiester bonds between two or more polynucleotides, which most often are double-stranded DNAs. Techniques for ligation are well known to the art and protocols for ligation are described in standard laboratory manuals and references, such as, *e.g.*, Sambrook (1989), *supra*.

Genome-derived "single exon probes," are probes that comprise at least part of an exon ("reference exon") and can hybridize detectably under high stringency conditions to transcript-derived nucleic acids that include the reference exon but do not hybridize detectably under high stringency conditions to nucleic acids that lack the reference exon. Single exon probes typically further comprise, contiguous to a first end of the exon portion, a first intronic and/or intergenic sequence that is identically contiguous to the exon in the genome, and may contain a second intronic and/or intergenic sequence that is identically contiguous to the exon in the genome. The minimum length of genome-derived single exon probes is defined by the requirement that the exonic portion be of sufficient length to hybridize under high stringency conditions to transcript-derived nucleic acids, as discussed above. The maximum length of genome-derived single exon probes is defined by the requirement that the probes contain portions of no more than one exon. The single exon probes may contain priming sequences not found in contiguity with the rest of the probe sequence in the genome, which priming sequences are useful for PCR and other amplification-based technologies. In another aspect, the invention is directed to single exon probes based on the CSNAs disclosed herein.

30 In one embodiment, the term "microarray" refers to a "nucleic acid microarray" having a substrate-bound plurality of nucleic acids, hybridization to each of the plurality of bound nucleic acids being separately detectable. The substrate can be solid or porous, planar or non-planar, unitary or distributed. Nucleic acid microarrays include all the

devices so called in Schena (ed.), DNA Microarrays: A Practical Approach (Practical Approach Series), Oxford University Press (1999); *Nature Genet.* 21(1)(suppl.):1 - 60 (1999); Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000). Additionally, these nucleic acid

5 microarrays include a substrate-bound plurality of nucleic acids in which the plurality of nucleic acids are disposed on a plurality of beads, rather than on a unitary planar substrate, as is described, *inter alia*, in Brenner *et al.*, *Proc. Natl. Acad. Sci. USA* 97(4):1665-1670 (2000). Examples of nucleic acid microarrays may be found in U.S. Patent Nos. 6,391,623, 6,383,754, 6,383,749, 6,380,377, 6,379,897, 6,376,191, 6,372,431, 6,351,712
10 6,344,316, 6,316,193, 6,312,906, 6,309,828, 6,309,824, 6,306,643, 6,300,063, 6,287,850, 6,284,497, 6,284,465, 6,280,954, 6,262,216, 6,251,601, 6,245,518, 6,263,287, 6,251,601, 6,238,866, 6,228,575, 6,214,587, 6,203,989, 6,171,797, 6,103,474, 6,083,726, 6,054,274, 6,040,138, 6,083,726, 6,004,755, 6,001,309, 5,958,342, 5,952,180, 5,936,731, 5,843,655, 5,814,454, 5,837,196, 5,436,327, 5,412,087, and 5,405,783, the disclosures of which are
15 incorporated herein by reference in their entireties.

In an alternative embodiment, a "microarray" may also refer to a "peptide microarray" or "protein microarray" having a substrate-bound collection or plurality of polypeptides, the binding to each of the plurality of bound polypeptides being separately detectable. Alternatively, the peptide microarray may have a plurality of binders,
20 including but not limited to monoclonal antibodies, polyclonal antibodies, phage display binders, yeast 2 hybrid binders, and aptamers, which can specifically detect the binding of the polypeptides of this invention. The array may be based on autoantibody detection to the polypeptides of this invention, see Robinson *et al.*, *Nature Medicine* 8(3):295-301 (2002). Examples of peptide arrays may be found in WO 02/31463, WO 02/25288, WO
25 01/94946, WO 01/88162, WO 01/68671, WO 01/57259, WO 00/61806, WO 00/54046, WO 00/47774, WO 99/40434, WO 99/39210, and WO 97/42507 and U.S. Patent Nos. 6,268,210, 5,766,960, and 5,143,854, the disclosures of which are incorporated herein by reference in their entireties.

In addition, determination of the levels of the CSNA or CSP may be made in a
30 multiplex manner using techniques described in WO 02/29109, WO 02/24959, WO 01/83502, WO01/73113, WO 01/59432, WO 01/57269, and WO 99/67641, the disclosures of which are incorporated herein by reference in their entireties.

The term “mutant”, “mutated”, or “mutation” when applied to nucleic acid sequences means that nucleotides in a nucleic acid sequence may be inserted, deleted or changed compared to a reference nucleic acid sequence. A single alteration may be made at a locus (a point mutation) or multiple nucleotides may be inserted, deleted or changed at a single locus. In addition, one or more alterations may be made at any number of loci within a nucleic acid sequence. In a preferred embodiment of the present invention, the nucleic acid sequence is the wild type nucleic acid sequence encoding a CSP or is a CSNA. The nucleic acid sequence may be mutated by any method known in the art including those mutagenesis techniques described *infra*.

The term “error-prone PCR” refers to a process for performing PCR under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. *See, e.g., Leung et al., Technique 1: 11-15 (1989) and Caldwell et al., PCR Methods Applic. 2: 28-33 (1992).*

The term “oligonucleotide-directed mutagenesis” refers to a process which enables the generation of site-specific mutations in any cloned DNA segment of interest. *See, e.g., Reidhaar-Olson et al., Science 241: 53-57 (1988).*

The term “assembly PCR” refers to a process which involves the assembly of a PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction.

The term “sexual PCR mutagenesis” or “DNA shuffling” refers to a method of error-prone PCR coupled with forced homologous recombination between DNA molecules of different but highly related DNA sequence *in vitro*, caused by random fragmentation of the DNA molecule based on sequence similarity, followed by fixation of the crossover by primer extension in an error-prone PCR reaction. *See, e.g., Stemmer, Proc. Natl. Acad. Sci. U.S.A. 91: 10747-10751 (1994).* DNA shuffling can be carried out between several related genes (“Family shuffling”).

The term “*in vivo* mutagenesis” refers to a process of generating random mutations in any cloned DNA of interest which involves the propagation of the DNA in a strain of bacteria such as *E. coli* that carries mutations in one or more of the DNA repair pathways. These “mutator” strains have a higher random mutation rate than that of a wild-type

parent. Propagating the DNA in a mutator strain will eventually generate random mutations within the DNA.

The term "cassette mutagenesis" refers to any process for replacing a small region of a double-stranded DNA molecule with a synthetic oligonucleotide "cassette" that
5 differs from the native sequence. The oligonucleotide often contains completely and/or partially randomized native sequence.

The term "recursive ensemble mutagenesis" refers to an algorithm for protein engineering (protein mutagenesis) developed to produce diverse populations of phenotypically related mutants whose members differ in amino acid sequence. This
10 method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. *See, e.g., Arkin et al., Proc. Natl. Acad. Sci. U.S.A.* 89: 7811-7815 (1992).

The term "exponential ensemble mutagenesis" refers to a process for generating combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position,
15 amino acids which lead to functional proteins. *See, e.g., Delegrave et al., Biotechnology Research* 11: 1548-1552 (1993); Arnold, *Current Opinion in Biotechnology* 4: 450-455 (1993).

"Operatively linked" expression control sequences refers to a linkage in which the expression control sequence is either contiguous with the gene of interest to control the
20 gene of interest, or acts in *trans* or at a distance to control the gene of interest.

The term "expression control sequence" as used herein refers to polynucleotide sequences which are necessary to affect the expression of coding sequences to which they are operatively linked. Expression control sequences are sequences which control the transcription, post-transcriptional events and translation of nucleic acid sequences.
25 Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*e.g.,* ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature
30 of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription termination sequence. The term "control sequences" is intended to include, at a minimum, all components whose presence is essential for expression, and can also include additional

components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

The term “vector,” as used herein, is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a circular double-stranded DNA loop into which additional DNA segments may be ligated. Other vectors include cosmids, bacterial artificial chromosomes (BAC) and yeast artificial chromosomes (YAC). Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Viral vectors that infect bacterial cells are referred to as bacteriophages. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “recombinant expression vectors” (or simply, “expression vectors”). In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, “plasmid” and “vector” may be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include other forms of expression vectors that serve equivalent functions.

The term “recombinant host cell” (or simply “host cell”), as used herein, is intended to refer to a cell into which a recombinant expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term “host cell” as used herein.

As used herein, the phrase “open reading frame” and the equivalent acronym “ORF” refers to that portion of a transcript-derived nucleic acid that can be translated in its entirety into a sequence of contiguous amino acids. As so defined, an ORF has length, measured in nucleotides, exactly divisible by 3. As so defined, an ORF need not encode the entirety of a natural protein.

As used herein, the phrase "ORF-encoded peptide" refers to the predicted or actual translation of an ORF.

As used herein, the phrase "degenerate variant" of a reference nucleic acid sequence is meant to be inclusive of all nucleic acid sequences that can be directly
5 translated, using the standard genetic code, to provide an amino acid sequence identical to that translated from the reference nucleic acid sequence.

The term "polypeptide" encompasses both naturally occurring and non-naturally occurring proteins and polypeptides, as well as polypeptide fragments and polypeptide mutants, derivatives and analogs thereof. A polypeptide may be monomeric or polymeric.
10 Further, a polypeptide may comprise a number of different modules within a single polypeptide each of which has one or more distinct activities. A preferred polypeptide in accordance with the invention comprises a CSP encoded by a nucleic acid molecule of the instant invention, or a fragment, mutant, analog or derivative thereof.

The term "isolated protein" or "isolated polypeptide" is a protein or polypeptide
15 that by virtue of its origin or source of derivation (1) is not associated with naturally associated components that accompany it in its native state, (2) is free of other proteins from the same species (3) is expressed by a cell from a different species, or (4) does not occur in nature. Thus, a polypeptide that is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be "isolated"
20 from its naturally associated components. A polypeptide or protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art.

A protein or polypeptide is "substantially pure," "substantially homogeneous" or "substantially purified" when at least about 60% to 75% of a sample exhibits a single
25 species of polypeptide. The polypeptide or protein may be monomeric or multimeric. A substantially pure polypeptide or protein will typically comprise about 50%, 60%, 70%, 80% or 90% W/W of a protein sample, more usually about 95%, and preferably will be over 99% pure. Protein purity or homogeneity may be determined by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample,
30 followed by visualizing a single polypeptide band upon staining the gel with a stain well known in the art. For certain purposes, higher resolution may be provided by using HPLC or other means well known in the art for purification.

The term "fragment" when used herein with respect to polypeptides of the present invention refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion compared to a full-length CSP. In a preferred embodiment, the fragment is a contiguous sequence in which the amino acid sequence of the fragment is identical to the corresponding positions in the naturally occurring polypeptide. Fragments typically are at least 5, 6, 7, 8, 9 or 10 amino acids long, preferably at least 12, 14, 16 or 18 amino acids long, more preferably at least 20 amino acids long, more preferably at least 25, 30, 35, 40 or 45, amino acids, even more preferably at least 50 or 60 amino acids long, and even more preferably at least 70 amino acids long.

A "derivative" when used herein with respect to polypeptides of the present invention refers to a polypeptide which is substantially similar in primary structural sequence to a CSP but which include, *e.g.*, *in vivo* or *in vitro* chemical and biochemical modifications that are not found in the CSP. Such modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Other modifications include, *e.g.*, labeling with radionuclides, and various enzymatic modifications, as will be readily appreciated by those skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes are well known in the art, and include radioactive isotopes such as ^{125}I , ^{32}P , ^{35}S , ^{14}C and ^3H , ligands which bind to labeled antiligands (*e.g.*, antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands which can serve as specific binding pair members for a labeled ligand. The choice of label depends on the sensitivity required, ease of conjugation with the primer, stability requirements, and available instrumentation. Methods for labeling polypeptides are well known in the art. *See* Ausubel (1992), *supra*; Ausubel (1999), *supra*.

The term "fusion protein" refers to polypeptides of the present invention coupled to a heterologous amino acid sequence. Fusion proteins are useful because they can be constructed to contain two or more desired functional elements from two or more different proteins. A fusion protein comprises at least 10 contiguous amino acids from a polypeptide of interest, more preferably at least 20 or 30 amino acids, even more preferably at least 40, 50 or 60 amino acids, yet more preferably at least 75, 100 or 125 amino acids. Fusion proteins can be produced recombinantly by constructing a nucleic acid sequence that encodes the polypeptide or a fragment thereof in frame with a nucleic acid sequence encoding a different protein or peptide and then expressing the fusion protein. Alternatively, a fusion protein can be produced chemically by crosslinking the polypeptide or a fragment thereof to another protein.

The term "analog" refers to both polypeptide analogs and non-peptide analogs. The term "polypeptide analog" as used herein refers to a polypeptide that is comprised of a segment of at least 25 amino acids that has substantial identity to a portion of an amino acid sequence but which contains non-natural amino acids or non-natural inter-residue bonds. In a preferred embodiment, the analog has the same or similar biological activity as the native polypeptide. Typically, polypeptide analogs comprise a conservative amino acid substitution (or insertion or deletion) with respect to the naturally occurring sequence. Analogs typically are at least 20 amino acids long, preferably at least 50 amino acids long or longer, and can often be as long as a full-length naturally occurring polypeptide.

The term "non-peptide analog" refers to a compound with properties that are analogous to those of a reference polypeptide. A non-peptide compound may also be termed a "peptide mimetic" or a "peptidomimetic." Such compounds are often developed with the aid of computerized molecular modeling. Peptide mimetics that are structurally similar to useful peptides may be used to produce an equivalent effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (*i.e.*, a polypeptide that has a desired biochemical property or pharmacological activity), but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of: $-\text{CH}_2\text{NH}-$, $-\text{CH}_2\text{S}-$, $-\text{CH}_2-\text{CH}_2-$, $-\text{CH}=\text{CH}-$ (cis and trans), $-\text{COCH}_2-$, $-\text{CH}(\text{OH})\text{CH}_2-$, and $-\text{CH}_2\text{SO}-$, by methods well known in the art. Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (*e.g.*, D-lysine in place of L-lysine) may also be used to generate more

stable peptides. In addition, constrained peptides comprising a consensus sequence or a substantially identical consensus sequence variation may be generated by methods known in the art (Rizo *et al.*, *Ann. Rev. Biochem.* 61:387-418 (1992)). For example, one may add internal cysteine residues capable of forming intramolecular disulfide bridges which
5 cyclize the peptide.

The term "mutant" or "mutein" when referring to a polypeptide of the present invention relates to an amino acid sequence containing substitutions, insertions or deletions of one or more amino acids compared to the amino acid sequence of a CSP. A mutein may have one or more amino acid point substitutions, in which a single amino acid
10 at a position has been changed to another amino acid, one or more insertions and/or deletions, in which one or more amino acids are inserted or deleted, respectively, in the sequence of the naturally occurring protein, and/or truncations of the amino acid sequence at either or both the amino or carboxy termini. Further, a mutein may have the same or different biological activity as the naturally occurring protein. For instance, a mutein may
15 have an increased or decreased biological activity. A mutein has at least 50% sequence similarity to the wild type protein, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are muteins having 80%, 85% or 90% sequence similarity to a CSP. In an even more preferred embodiment, a mutein exhibits 95% sequence identity, even more preferably 97%, even more preferably 98% and even
20 more preferably 99%. Sequence similarity may be measured by any common sequence analysis algorithm, such as GAP or BESTFIT or other variation Smith-Waterman alignment. *See*, T. F. Smith and M. S. Waterman, *J. Mol. Biol.* 147:195-197 (1981) and W.R. Pearson, *Genomics* 11:635-650 (1991).

Preferred amino acid substitutions are those which: (1) reduce susceptibility to
25 proteolysis, (2) reduce susceptibility to oxidation, (3) alter binding affinity for forming protein complexes, (4) alter binding affinity or enzymatic activity, and (5) confer or modify other physicochemical or functional properties of such analogs. For example, single or multiple amino acid substitutions (preferably conservative amino acid substitutions) may be made in the naturally occurring sequence (preferably in the portion
30 of the polypeptide outside the domain(s) forming intermolecular contacts. In a preferred embodiment, the amino acid substitutions are moderately conservative substitutions or conservative substitutions. In a more preferred embodiment, the amino acid substitutions are conservative substitutions. A conservative amino acid substitution should not

substantially change the structural characteristics of the parent sequence (*e.g.*, a replacement amino acid should not tend to disrupt a helix that occurs in the parent sequence, or disrupt other types of secondary structure that characterize the parent sequence). Examples of art-recognized polypeptide secondary and tertiary structures are described in Creighton (ed.), Proteins, Structures and Molecular Principles, W. H. Freeman and Company (1984); Branden *et al.* (ed.), Introduction to Protein Structure, Garland Publishing (1991); Thornton *et al.*, *Nature* 354:105-106 (1991).

As used herein, the twenty conventional amino acids and their abbreviations follow conventional usage. See Golub *et al.* (eds.), Immunology - A Synthesis 2nd Ed., Sinauer Associates (1991). Stereoisomers (*e.g.*, D-amino acids) of the twenty conventional amino acids, unnatural amino acids such as α -, α -disubstituted amino acids, N-alkyl amino acids, and other unconventional amino acids may also be suitable components for polypeptides of the present invention. Examples of unconventional amino acids include: 4-hydroxyproline, γ -carboxyglutamate, ϵ -N,N,N-trimethyllysine, ϵ -N-acetyllysine, O-phosphoserine, N-acetylserine, N-formylmethionine, 3-methylhistidine, 5-hydroxylysine, *s*-N-methylarginine, and other similar amino acids and imino acids (*e.g.*, 4-hydroxyproline). In the polypeptide notation used herein, the left-hand direction is the amino terminal direction and the right hand direction is the carboxy-terminal direction, in accordance with standard usage and convention.

By "homology" or "homologous" when referring to a polypeptide of the present invention it is meant polypeptides from different organisms with a similar sequence to the encoded amino acid sequence of a CSP and a similar biological activity or function. Although two polypeptides are said to be "homologous," this does not imply that there is necessarily an evolutionary relationship between the polypeptides. Instead, the term "homologous" is defined to mean that the two polypeptides have similar amino acid sequences and similar biological activities or functions. In a preferred embodiment, a homologous polypeptide is one that exhibits 50% sequence similarity to CSP, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are homologous polypeptides that exhibit 80%, 85% or 90% sequence similarity to a CSP. In yet a more preferred embodiment, a homologous polypeptide exhibits 95%, 97%, 98% or 99% sequence similarity.

When "sequence similarity" is used in reference to polypeptides, it is recognized that residue positions that are not identical often differ by conservative amino acid

substitutions. In a preferred embodiment, a polypeptide that has "sequence similarity" comprises conservative or moderately conservative amino acid substitutions. A "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (*e.g.*, charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative substitutions, the percent sequence identity or degree of similarity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. *See, e.g.*, Pearson, *Methods Mol. Biol.* 24: 307-31 (1994).

For instance, the following six groups each contain amino acids that are conservative substitutions for one another:

- 1) Serine (S), Threonine (T);
- 2) Aspartic Acid (D), Glutamic Acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Alternatively, a conservative replacement is any change having a positive value in the PAM250 log-likelihood matrix disclosed in Gonnet *et al.*, *Science* 256: 1443-45 (1992). A "moderately conservative" replacement is any change having a nonnegative value in the PAM250 log-likelihood matrix.

Sequence similarity for polypeptides, which is also referred to as sequence identity, is typically measured using sequence analysis software. Protein analysis software matches similar sequences using measures of similarity assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild type protein and a mutein thereof. *See, e.g.*, GCG Version 6.1. Other programs include FASTA, discussed *supra*.

A preferred algorithm when comparing a sequence of the invention to a database containing a large number of sequences from different organisms is the computer program BLAST, especially blastp or tblastn. *See, e.g., Altschul et al., J. Mol. Biol.* 215: 403-410 (1990); Altschul et al., *Nucleic Acids Res.* 25:3389-402 (1997). Preferred parameters for

5 blastp are:

	Expectation value:	10 (default)
	Filter:	seg (default)
	Cost to open a gap:	11 (default)
	Cost to extend a gap:	1 (default)
10	Max. alignments:	100 (default)
	Word size:	11 (default)
	No. of descriptions:	100 (default)
	Penalty Matrix:	BLOSUM62

The length of polypeptide sequences compared for homology will generally be at
15 least about 16 amino acid residues, usually at least about 20 residues, more usually at least about 24 residues, typically at least about 28 residues, and preferably more than about 35 residues. When searching a database containing sequences from a large number of different organisms, it is preferable to compare amino acid sequences.

Algorithms other than blastp for database searching using amino acid sequences
20 are known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA (*e.g.*, FASTA2 and FASTA3) provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson (1990), *supra*; Pearson (2000), *supra*. For example, percent sequence identity between amino acid sequences can be determined using FASTA
25 with its default or recommended parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1.

An "antibody" refers to an intact immunoglobulin, or to an antigen-binding portion thereof that competes with the intact antibody for specific binding to a molecular species, *e.g.*, a polypeptide of the instant invention. Antigen-binding portions may be produced by
30 recombinant DNA techniques or by enzymatic or chemical cleavage of intact antibodies. Antigen-binding portions include, *inter alia*, Fab, Fab', F(ab')₂, Fv, dAb, and complementarity determining region (CDR) fragments, single-chain antibodies (scFv), chimeric antibodies, diabodies and polypeptides that contain at least a portion of an

immunoglobulin that is sufficient to confer specific antigen binding to the polypeptide. A Fab fragment is a monovalent fragment consisting of the VL, VH, CL and CH1 domains; a F(ab')₂ fragment is a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; a Fd fragment consists of the VH and CH1 domains; a Fv fragment consists of the VL and VH domains of a single arm of an antibody; and a dAb fragment consists of a VH domain. See, e.g., Ward *et al.*, *Nature* 341: 544-546 (1989).

By "bind specifically" and "specific binding" as used herein it is meant the ability of the antibody to bind to a first molecular species in preference to binding to other molecular species with which the antibody and first molecular species are admixed. An antibody is said to "recognize" a first molecular species when it can bind specifically to that first molecular species.

A single-chain antibody (scFv) is an antibody in which VL and VH regions are paired to form a monovalent molecule via a synthetic linker that enables them to be made as a single protein chain. See, e.g., Bird *et al.*, *Science* 242: 423-426 (1988); Huston *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 5879-5883 (1988). Diabodies are bivalent, bispecific antibodies in which VH and VL domains are expressed on a single polypeptide chain, but using a linker that is too short to allow for pairing between the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen binding sites. See e.g., Holliger *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 6444-6448 (1993); Poljak *et al.*, *Structure* 2: 1121-1123 (1994). One or more CDRs may be incorporated into a molecule either covalently or noncovalently to make it an immunoadhesin. An immunoadhesin may incorporate the CDR(s) as part of a larger polypeptide chain, may covalently link the CDR(s) to another polypeptide chain, or may incorporate the CDR(s) noncovalently. The CDRs permit the immunoadhesin to specifically bind to a particular antigen of interest. A chimeric antibody is an antibody that contains one or more regions from one antibody and one or more regions from one or more other antibodies.

An antibody may have one or more binding sites. If there is more than one binding site, the binding sites may be identical to one another or may be different. For instance, a naturally occurring immunoglobulin has two identical binding sites, a single-chain antibody or Fab fragment has one binding site, while a "bispecific" or "bifunctional" antibody has two different binding sites.

An "isolated antibody" is an antibody that (1) is not associated with naturally-associated components, including other naturally-associated antibodies, that accompany it in its native state, (2) is free of other proteins from the same species, (3) is expressed by a cell from a different species, or (4) does not occur in nature. It is known that purified
5 proteins, including purified antibodies, may be stabilized with non-naturally-associated components. The non-naturally-associated component may be a protein, such as albumin (e.g., BSA) or a chemical such as polyethylene glycol (PEG).

A "neutralizing antibody" or "an inhibitory antibody" is an antibody that inhibits the activity of a polypeptide or blocks the binding of a polypeptide to a ligand that
10 normally binds to it. An "activating antibody" is an antibody that increases the activity of a polypeptide.

The term "epitope" includes any protein determinant capable of specific binding to an immunoglobulin or T-cell receptor. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains
15 and usually have specific three-dimensional structural characteristics, as well as specific charge characteristics. An antibody is said to specifically bind an antigen when the dissociation constant is less than $1\ \mu\text{M}$, preferably less than $100\ \text{nM}$ and most preferably less than $10\ \text{nM}$.

The term "patient" includes human and veterinary subjects.
20 Throughout this specification and claims, the word "comprise," or variations such as "comprises" or "comprising," will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

The term "colon specific" refers to a nucleic acid molecule or polypeptide that is expressed predominantly in the colon as compared to other tissues in the body. In a
25 preferred embodiment, a "colon specific" nucleic acid molecule or polypeptide is detected at a level that is 1.5-fold higher than any other tissue in the body. In a more preferred embodiment, the "colon specific" nucleic acid molecule or polypeptide is detected at a level that is 2-fold higher than any other tissue in the body, more preferably 5-fold higher, still more preferably at least 10-fold, 15-fold, 20-fold, 25-fold, 50-fold or 100-fold higher
30 than any other tissue in the body. Nucleic acid molecule levels may be measured by nucleic acid hybridization, such as Northern blot hybridization, or quantitative PCR. Polypeptide levels may be measured by any method known to accurately quantitate protein levels, such as Western blot analysis.

Nucleic Acid Molecules, Regulatory Sequences, Vectors, Host Cells and Recombinant Methods of Making Polypeptides

Nucleic Acid Molecules

One aspect of the invention provides isolated nucleic acid molecules that are
5 specific to the colon or to colon cells or tissue or that are derived from such nucleic acid molecules. These isolated colon specific nucleic acids (CSNAs) may comprise cDNA genomic DNA, RNA, or a combination thereof, a fragment of one of these nucleic acids, or may be a non-naturally occurring nucleic acid molecule. A CSNA may be derived from an animal. In a preferred embodiment, the CSNA is derived from a human or other
10 mammal. In a more preferred embodiment, the CSNA is derived from a human or other primate. In an even more preferred embodiment, the CSNA is derived from a human.

In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that is specific to colon, a colon-specific polypeptide (CSP). In a more preferred embodiment, the nucleic acid molecule encodes a polypeptide that comprises an amino acid sequence of
15 SEQ ID NO: 113-259. In another highly preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-112. Nucleotide sequences of the instantly-described nucleic acid molecules were determined by assembling several DNA molecules from either public or proprietary databases. Some of the underlying DNA sequences are the result, directly or indirectly, of at least one enzymatic polymerization
20 reaction (*e.g.*, reverse transcription and/or polymerase chain reaction) using an automated sequencer (such as the MegaBACE™ 1000, Amersham Biosciences, Sunnyvale, CA, USA).

Nucleic acid molecules of the present invention may also comprise sequences that selectively hybridize to a nucleic acid molecule encoding a CSNA or a complement or
25 antisense thereof. The hybridizing nucleic acid molecule may or may not encode a polypeptide or may or may not encode a CSP. However, in a preferred embodiment, the hybridizing nucleic acid molecule encodes a CSP. In a more preferred embodiment, the invention provides a nucleic acid molecule that selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule that encodes a polypeptide
30 comprising an amino acid sequence of SEQ ID NO: 113-259. In an even more preferred embodiment, the invention provides a nucleic acid molecule that selectively hybridizes to a nucleic acid molecule comprising the nucleic acid sequence of SEQ ID NO: 1-112 or the antisense sequence thereof. Preferably, the nucleic acid molecule selectively hybridizes to

a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a CSP under low stringency conditions. More preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a CSP under moderate stringency conditions. Most preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a CSP under high stringency conditions. In a preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high stringency conditions to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence of SEQ ID NO: 113-259. In a more preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high stringency conditions to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule comprising a nucleic acid sequence selected from SEQ ID NO: 1-112.

Nucleic acid molecules of the present invention may also comprise nucleic acid sequences that exhibit substantial sequence similarity to a nucleic acid encoding a CSP or a complement of the encoding nucleic acid molecule. In this embodiment, it is preferred that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid molecule encoding human CSP. More preferred is a nucleic acid molecule exhibiting substantial sequence similarity to a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 113-259. By substantial sequence similarity it is meant a nucleic acid molecule having at least 60%, more preferably at least 70%, even more preferably at least 80% and even more preferably at least 85% sequence identity with a nucleic acid molecule encoding a CSP, such as a polypeptide having an amino acid sequence of SEQ ID NO: 113-259. In a more preferred embodiment, the similar nucleic acid molecule is one that has at least 90%, more preferably at least 95%, more preferably at least 97%, even more preferably at least 98%, and still more preferably at least 99% sequence identity with a nucleic acid molecule encoding a CSP. Most preferred in this embodiment is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a nucleic acid molecule encoding a CSP.

The nucleic acid molecules of the present invention are also inclusive of those exhibiting substantial sequence similarity to a CSNA or its complement. In this embodiment, it is preferred that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid molecule having a nucleic acid sequence of SEQ ID NO: 1-

112. By substantial sequence similarity it is meant a nucleic acid molecule that has at least 60%, more preferably at least 70%, even more preferably at least 80% and even more preferably at least 85% sequence identity with a CSNA, such as one having a nucleic acid sequence of SEQ ID NO: 1-112. More preferred is a nucleic acid molecule that has at least 90%, more preferably at least 95%, more preferably at least 97%, even more preferably at least 98%, and still more preferably at least 99% sequence identity with a CSNA. Most preferred is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a CSNA.

Nucleic acid molecules that exhibit substantial sequence similarity are inclusive of sequences that exhibit sequence identity over their entire length to a CSNA or to a nucleic acid molecule encoding a CSP, as well as sequences that are similar over only a part of its length. In this case, the part is at least 50 nucleotides of the CSNA or the nucleic acid molecule encoding a CSP, preferably at least 100 nucleotides, more preferably at least 150 or 200 nucleotides, even more preferably at least 250 or 300 nucleotides, still more preferably at least 400 or 500 nucleotides.

The substantially similar nucleic acid molecule may be a naturally occurring one that is derived from another species, especially one derived from another primate, wherein the similar nucleic acid molecule encodes an amino acid sequence that exhibits significant sequence identity to that of SEQ ID NO: 113-259 or demonstrates significant sequence identity to the nucleotide sequence of SEQ ID NO: 1-112. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule from a human, when the CSNA is a member of a gene family. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-primate, mammalian species, including without limitation, domesticated species, *e.g.*, dog, cat, mouse, rat, rabbit, hamster, cow, horse and pig; and wild animals, *e.g.*, monkey, fox, lions, tigers, bears, giraffes, zebras, etc. The substantially similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-mammalian species, such as birds or reptiles. The naturally occurring substantially similar nucleic acid molecule may be isolated directly from humans or other species. In another embodiment, the substantially similar nucleic acid molecule may be one that is experimentally produced by random mutation of a nucleic acid molecule. In another embodiment, the substantially similar nucleic acid molecule may be one that is experimentally produced by directed

mutation of a CSNA. In a preferred embodiment, the substantially similar nucleic acid molecule is a CSNA.

The nucleic acid molecules of the present invention are also inclusive of allelic variants of a CSNA or a nucleic acid encoding a CSP. For example, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes and the sequence
5 determined from one individual of a species may differ from other allelic forms present within the population. More than 1.4 million SNPs have already been identified in the human genome, International Human Genome Sequencing Consortium, *Nature* 409: 860-921 (2001) – Variants with small deletions and insertions of more than a single nucleotide
10 are also found in the general population, and often do not alter the function of the protein. In addition, amino acid substitutions occur frequently among natural allelic variants, and often do not substantially change protein function.

In a preferred embodiment, the allelic variant is a variant of a gene, wherein the gene is transcribed into a mRNA that encodes a CSP. In a more preferred embodiment,
15 the gene is transcribed into a mRNA that encodes a CSP comprising an amino acid sequence of SEQ ID NO: 113-259. In another preferred embodiment, the allelic variant is a variant of a gene, wherein the gene is transcribed into a mRNA that is a CSNA. In a more preferred embodiment, the gene is transcribed into a mRNA that comprises the nucleic acid sequence of SEQ ID NO: 1-112. Also preferred is that the allelic variant be a
20 naturally occurring allelic variant in the species of interest, particularly human.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences comprising a part of a nucleic acid sequence of the instant invention. The part may or may not encode a polypeptide, and may or may not encode a polypeptide that is a CSP. In a preferred embodiment, the part encodes a CSP. In one embodiment, the
25 nucleic acid molecule comprises a part of a CSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that hybridizes or exhibits substantial sequence similarity to a CSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that is an allelic variant of a CSNA. In yet another embodiment, the nucleic acid molecule comprises a part of a nucleic acid
30 molecule that encodes a CSP. A part comprises at least 10 nucleotides, more preferably at least 15, 17, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides. The maximum size of a nucleic acid part is one nucleotide shorter than the sequence of the nucleic acid molecule encoding the full-length protein.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences that encode fusion proteins, homologous proteins, polypeptide fragments, muteins and polypeptide analogs, as described *infra*.

5 Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences containing modifications of the native nucleic acid molecule. Examples of such modifications include, but are not limited to, nonnative internucleoside bonds, post-synthetic modifications or altered nucleotide analogues. One having ordinary skill in the art would recognize that the type of modification that may be made will depend upon the intended use of the nucleic acid molecule. For instance, when the nucleic acid molecule is
10 used as a hybridization probe, the range of such modifications will be limited to those that permit sequence-discriminating base pairing of the resulting nucleic acid. When used to direct expression of RNA or protein *in vitro* or *in vivo*, the range of such modifications will be limited to those that permit the nucleic acid to function properly as a polymerization substrate. When the isolated nucleic acid is used as a therapeutic agent,
15 the modifications will be limited to those that do not confer toxicity upon the isolated nucleic acid.

Accordingly, in one embodiment, a nucleic acid molecule may include nucleotide analogues that incorporate labels that are directly detectable, such as radiolabels or fluorophores, or nucleotide analogues that incorporate labels that can be visualized in a
20 subsequent reaction, such as biotin or various haptens. The labeled nucleic acid molecules are particularly useful as hybridization probes.

Common radiolabeled analogues include those labeled with ^{33}P , ^{32}P , and ^{35}S , such as α - ^{32}P -dATP, α - ^{32}P -dCTP, α - ^{32}P -dGTP, α - ^{32}P -dTTP, α - ^{32}P -3'dATP, α - ^{32}P -ATP, α - ^{32}P -CTP, α - ^{32}P -GTP, α - ^{32}P -UTP, α - ^{35}S -dATP, γ - ^{35}S -GTP, γ - ^{33}P -dATP, and the like.

25 Commercially available fluorescent nucleotide analogues readily incorporated into the nucleic acids of the present invention include Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy3-dUTP (Amersham Biosciences, Piscataway, New Jersey, USA), fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP, Texas Red®-5-dUTP, Cascade Blue®-7-dUTP, BODIPY® FL-14-dUTP, BODIPY® TMR-14-dUTP, BODIPY® TR-14-dUTP,
30 Rhodamine Green™-5-dUTP, Oregon Green® 488-5-dUTP, Texas Red®-12-dUTP, BODIPY® 630/650-14-dUTP, BODIPY® 650/665-14-dUTP, Alexa Fluor® 488-5-dUTP, Alexa Fluor® 532-5-dUTP, Alexa Fluor® 568-5-dUTP, Alexa Fluor® 594-5-dUTP, Alexa Fluor® 546-14-dUTP, fluorescein-12-UTP, tetramethylrhodamine-6-UTP, Texas

Red®-5-UTP, Cascade Blue®-7-UTP, BODIPY® FL-14-UTP, BODIPY® TMR-14-UTP, BODIPY® TR-14-UTP, Rhodamine Green™-5-UTP, Alexa Fluor® 488-5-UTP, Alexa Fluor® 546-14-UTP (Molecular Probes, Inc. Eugene, OR, USA). One may also custom synthesize nucleotides having other fluorophores. See Henegariu *et al.*, *Nature*

5 *Biotechnol.* 18: 345-348 (2000).

Haptens that are commonly conjugated to nucleotides for subsequent labeling include biotin (biotin-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA; biotin-21-UTP, biotin-21-dUTP, Clontech Laboratories, Inc., Palo Alto, CA, USA), digoxigenin (DIG-11-dUTP, alkali labile, DIG-11-UTP, Roche Diagnostics Corp., Indianapolis, IN, USA), and dinitrophenyl (dinitrophenyl-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA).

Nucleic acid molecules of the present invention can be labeled by incorporation of labeled nucleotide analogues into the nucleic acid. Such analogues can be incorporated by enzymatic polymerization, such as by nick translation, random priming, polymerase chain reaction (PCR), terminal transferase tailing, and end-filling of overhangs, for DNA molecules, and *in vitro* transcription driven, *e.g.*, from phage promoters, such as T7, T3, and SP6, for RNA molecules. Commercial kits are readily available for each such labeling approach. Analogues can also be incorporated during automated solid phase chemical synthesis. Labels can also be incorporated after nucleic acid synthesis, with the 5' phosphate and 3' hydroxyl providing convenient sites for post-synthetic covalent attachment of detectable labels.

Other post-synthetic approaches also permit internal labeling of nucleic acids. For example, fluorophores can be attached using a cisplatin reagent that reacts with the N7 of guanine residues (and, to a lesser extent, adenine bases) in DNA, RNA, and Peptide Nucleic Acids (PNA) to provide a stable coordination complex between the nucleic acid and fluorophore label (Universal Linkage System) (available from Molecular Probes, Inc., Eugene, OR, USA and Amersham Pharmacia Biotech, Piscataway, NJ, USA); see Alers *et al.*, *Genes, Chromosomes & Cancer* 25: 301- 305 (1999); Jelsma *et al.*, *J. NIH Res.* 5: 82 (1994); Van Belkum *et al.*, *BioTechniques* 16: 148-153 (1994). Alternatively, nucleic acids can be labeled using a disulfide-containing linker (FastTag™ Reagent, Vector Laboratories, Inc., Burlingame, CA, USA) that is photo- or thermally coupled to the target nucleic acid using aryl azide chemistry; after reduction, a free thiol is available for coupling to a hapten, fluorophore, sugar, affinity ligand, or other marker.

One or more independent or interacting labels can be incorporated into the nucleic acid molecules of the present invention. For example, both a fluorophore and a moiety that in proximity thereto acts to quench fluorescence can be included to report specific hybridization through release of fluorescence quenching or to report exonucleotidic excision. See, e.g., Tyagi *et al.*, *Nature Biotechnol.* 14: 303-308 (1996); Tyagi *et al.*, *Nature Biotechnol.* 16: 49-53 (1998); Sokol *et al.*, *Proc. Natl. Acad. Sci. USA* 95: 11538-11543 (1998); Kostrikis *et al.*, *Science* 279: 1228-1229 (1998); Marras *et al.*, *Genet. Anal.* 14: 151-156 (1999); Holland *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 7276-7280 (1991); Heid *et al.*, *Genome Res.* 6(10): 986-94 (1996); Kuimelis *et al.*, *Nucleic Acids Symp. Ser.* (37): 255-6 (1997); and U.S. Patent Nos. 5,846,726, 5,925,517, 5,925,517, 5,723,591 and 5,538,848, the disclosures of which are incorporated herein by reference in their entirety.

Nucleic acid molecules of the present invention may also be modified by altering one or more native phosphodiester internucleoside bonds to more nuclease-resistant, internucleoside bonds. See Hartmann *et al.* (eds.), Manual of Antisense Methodology: Perspectives in Antisense Science, Kluwer Law International (1999); Stein *et al.* (eds.), Applied Antisense Oligonucleotide Technology, Wiley-Liss (1998); Chadwick *et al.* (eds.), Oligonucleotides as Therapeutic Agents – Symposium No. 209, John Wiley & Son Ltd (1997). Such altered internucleoside bonds are often desired for techniques or for targeted gene correction, Gamper *et al.*, *Nucl. Acids Res.* 28(21): 4332-4339 (2000). For double-stranded RNA inhibition which may utilize either natural ds RNA or ds RNA modified in its, sugar, phosphate or base, see Hannon, *Nature* 418(11): 244-251 (2002); Fire *et al.* in WO 99/32619; Tuschl *et al.* in US2002/0086356; Krutetz *et al.* in WO 00/44895, the disclosures of which are incorporated herein by reference in their entirety. For circular antisense, see Kool in U.S. Patent No. 5,426,180, the disclosure of which is incorporated herein by reference in its entirety.

Modified oligonucleotide backbones include, without limitation, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity

wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'.

Representative U.S. Patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S. Patent Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050, the disclosures of which are incorporated herein by reference in their entireties. In a preferred embodiment, the modified internucleoside linkages may be used for antisense techniques.

Other modified oligonucleotide backbones do not include a phosphorus atom, but have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts. Representative U.S. patents that teach the preparation of the above backbones include, but are not limited to, U.S. Patent Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437 and 5,677,439; the disclosures of which are incorporated herein by reference in their entireties.

In other preferred nucleic acid molecules, both the sugar and the internucleoside linkage are replaced with novel groups, such as peptide nucleic acids (PNA). In PNA compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing backbone, in particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. Nucleobases are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone, typically by methylene carbonyl linkages. PNA can be synthesized using a modified peptide synthesis protocol. PNA oligomers can be synthesized by both Fmoc and tBoc methods. Representative U.S. patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Patent Nos.

5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference in its entirety. Automated PNA synthesis is readily achievable on commercial synthesizers (see, e.g., "PNA User's Guide," Rev. 2, February 1998, Perseptive Biosystems Part No. 60138, Applied Biosystems, Inc., Foster City, CA). PNA molecules are advantageous for a number of reasons. First, because the PNA backbone is uncharged, PNA/DNA and PNA/RNA duplexes have a higher thermal stability than is found in DNA/DNA and DNA/RNA duplexes. The T_m of a PNA/DNA or PNA/RNA duplex is generally 1°C higher per base pair than the T_m of the corresponding DNA/DNA or DNA/RNA duplex (in 100 mM NaCl). Second, PNA molecules can also form stable PNA/DNA complexes at low ionic strength, under conditions in which DNA/DNA duplex formation does not occur. Third, PNA also demonstrates greater specificity in binding to complementary DNA because a PNA/DNA mismatch is more destabilizing than DNA/DNA mismatch. A single mismatch in mixed a PNA/DNA 15-mer lowers the T_m by 8–20°C (15°C on average). In the corresponding DNA/DNA duplexes, a single mismatch lowers the T_m by 4–16°C (11°C on average). Because PNA probes can be significantly shorter than DNA probes, their specificity is greater. Fourth, PNA oligomers are resistant to degradation by enzymes, and the lifetime of these compounds is extended both *in vivo* and *in vitro* because nucleases and proteases do not recognize the PNA polyamide backbone with nucleobase sidechains. See, e.g., Ray *et al.*, *FASEB J.* 14(9): 1041-60 (2000); Nielsen *et al.*, *Pharmacol Toxicol.* 86(1): 3-7 (2000); Larsen *et al.*, *Biochim Biophys Acta.* 1489(1): 159-66 (1999); Nielsen, *Curr. Opin. Struct. Biol.* 9(3): 353-7 (1999), and Nielsen, *Curr. Opin. Biotechnol.* 10(1): 71-5 (1999).

Nucleic acid molecules may be modified compared to their native structure throughout the length of the nucleic acid molecule or can be localized to discrete portions thereof. As an example of the latter, chimeric nucleic acids can be synthesized that have discrete DNA and RNA domains and that can be used for targeted gene repair and modified PCR reactions, as further described in, Misra *et al.*, *Biochem.* 37: 1917-1925 (1998); and Finn *et al.*, *Nucl. Acids Res.* 24: 3357-3363 (1996), and U.S. Patent Nos. 5,760,012 and 5,731,181, the disclosures of which are incorporated herein by reference in their entireties.

Unless otherwise specified, nucleic acid molecules of the present invention can include any topological conformation appropriate to the desired use; the term thus explicitly comprehends, among others, single-stranded, double-stranded, triplexed,

quadruplexed, partially double-stranded, partially-triplexed, partially-quadruplexed, branched, hairpinned, circular, and padlocked conformations. Padlocked conformations and their utilities are further described in Banér *et al.*, *Curr. Opin. Biotechnol.* 12: 11-15 (2001); Escude *et al.*, *Proc. Natl. Acad. Sci. USA* 14: 96(19):10603-7 (1999); and Nilsson
5 *et al.*, *Science* 265(5181): 2085-8 (1994). Triplexed and quadruplexed conformations, and their utilities, are reviewed in Praseuth *et al.*, *Biochim. Biophys. Acta.* 1489(1): 181-206 (1999); Fox, *Curr. Med. Chem.* 7(1): 17-37 (2000); Kochetkova *et al.*, *Methods Mol. Biol.* 130: 189-201 (2000); Chan *et al.*, *J. Mol. Med.* 75(4): 267-82 (1997); Rowley *et al.*, *Mol Med* 5(10): 693-700 (1999); Kool, *Annu Rev Biophys Biomol Struct.* 25: 1-28 (1996).

10 *SNP Polymorphisms*

Commonly, sequence differences between individuals involve differences in single nucleotide positions. SNPs may account for 90% of human DNA polymorphism. Collins
et al., 8 *Genome Res.* 1229-31 (1998). SNPs include single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in a population. In addition,
15 the least frequent allele generally must occur at a frequency of 1% or greater. DNA sequence variants with a reasonably high population frequency are observed approximately every 1,000 nucleotide across the genome, with estimates as high as 1 SNP per 350 base pairs. Wang *et al.*, 280 *Science* 1077-82 (1998); Harding *et al.*, 60 *Am. J. Human Genet.* 772-89 (1997); Taillon-Miller *et al.*, 8 *Genome Res.* 748-54 (1998); Cargill
20 *et al.*, 22 *Nat. Genet.* 231-38 (1999); and Semple *et al.*, 16 *Bioinform. Disc. Note* 735-38 (2000). The frequency of SNPs varies with the type and location of the change. In base substitutions, two-thirds of the substitutions involve the C-T and G-A type. This variation in frequency can be related to 5-methylcytosine deamination reactions that occur frequently, particularly at CpG dinucleotides. Regarding location, SNPs occur at a much
25 higher frequency in non-coding regions than in coding regions. Information on over one million variable sequences is already publicly available via the Internet and more such markers are available from commercial providers of genetic information. Kwok and Gu, 5 *Med. Today* 538-53 (1999).

Several definitions of SNPs exist. See, e.g., Brooks, 235 *Gene* 177-86 (1999). As
30 used herein, the term "single nucleotide polymorphism" or "SNP" includes all single base variants, thus including nucleotide insertions and deletions in addition to single nucleotide substitutions. There are two types of nucleotide substitutions. A transition is the

replacement of one purine by another purine or one pyrimidine by another pyrimidine. A transversion is the replacement of a purine for a pyrimidine, or vice versa.

Numerous methods exist for detecting SNPs within a nucleotide sequence. A review of many of these methods can be found in Landegren *et al.*, 8 *Genome Res.* 769-76
5 (1998). For example, a SNP in a genomic sample can be detected by preparing a Reduced Complexity Genome (RCG) from the genomic sample, then analyzing the RCG for the presence or absence of a SNP. See, e.g., WO 00/18960 which is herein incorporated by reference in its entirety. Multiple SNPs in a population of target polynucleotides in parallel can be detected using, for example, the methods of WO 00/50869 which is herein
10 incorporated by reference in its entirety. Other SNP detection methods include the methods of U.S. Pat. Nos. 6,297,018 and 6,322,980 which are herein incorporated by reference in their entirety. Furthermore, SNPs can be detected by restriction fragment length polymorphism (RFLP) analysis. See, e.g., U.S. Pat. Nos. 5,324,631; 5,645,995 which are herein incorporated by reference in their entirety. RFLP analysis of SNPs,
15 however, is limited to cases where the SNP either creates or destroys a restriction enzyme cleavage site. SNPs can also be detected by direct sequencing of the nucleotide sequence of interest. In addition, numerous assays based on hybridization have also been developed to detect SNPs and mismatch distinction by polymerases and ligases. Several web sites provide information about SNPs including Ensembl on the World Wide Web at
20 ensemble.org, Sanger Institute on the World Wide Web at sanger.ac.uk/genetics/exon/, National Center for Biotechnology Information (NCBI) on the World Wide Web at ncbi.nlm.nih.gov/SNP/, The SNP Consortium Ltd. on the World Wide Web at snp.cshl.org. The chromosomal locations for the compositions disclosed herein are provided below. In addition, one of ordinary skill in the art could use a BLAST against
25 the genome or any of the databases cited above to find the chromosomal location. Another a preferred method to find the genomic coordinates and associated SNPs would be to use the BLAT tool (genome.ucsc.edu, Kent et al. 2001, The Human Genome Browser at UCSC, Genome Research 996-1006 or Kent 2002 BLAT —The BLAST -Like Alignment Tool Genome Research, 1-9). All web sites above were accessed December 3,
30 2003.

RNA interference

RNA interference refers to the process of sequence-specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA). Fire *et al.*, 1998, *Nature*, 391, 806. The corresponding process in plants is commonly referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi.

5 The process of post transcriptional gene silencing is thought to be an evolutionarily conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla. Fire *et al.*, 1999, *Trends Genet.*, 15, 358. Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNA) derived from viral infection or the

10 random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response though a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-

15 oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA). Berstein *et al.*, 2001, *Nature*, 409, 363. Short interfering RNAs derived from dicer activity are typically about

20 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control. Hutvagner *et al.*, 2001, *Science*, 293, 834. The RNAi response also features an

25 endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex. Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188.

Short interfering RNA mediated RNAi has been studied in a variety of systems.

30 Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in *C. Elegans*. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describe RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in *Drosophila* cells

transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in *Drosophila* embryonic lysates (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877) has revealed certain requirements for
5 siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two nucleotide 3'-overhangs. Furthermore, complete substitution of one or both siRNA strands with 2'-deoxy (2'-H) or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of the 3'-terminal siRNA overhang
10 nucleotides with deoxy nucleotides (2'-H) was shown to be tolerated. Single mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end. Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877. Other studies have indicated that a 5'-phosphate on the
15 target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA. Nykanen *et al.*, 2001, *Cell*, 107, 309.

Studies have shown that replacing the 3'-overhanging segments of a 21-mer siRNA duplex having 2 nucleotide 3' overhangs with deoxyribonucleotides does not have an
20 adverse effect on RNAi activity. Replacing up to 4 nucleotides on each end of the siRNA with deoxyribonucleotides has been reported to be well tolerated whereas complete substitution with deoxyribonucleotides results in no RNAi activity. Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877. In addition, Elbashir *et al.*, *supra*, also report that substitution of siRNA with 2'-O-methyl nucleotides completely abolishes RNAi activity. Li *et al.*, WO
25 00/44914, and Beach *et al.*, WO 01/68836 both suggest that siRNA "may include modifications to either the phosphate-sugar back bone or the nucleoside to include at least one of a nitrogen or sulfur heteroatom", however neither application teaches to what extent these modifications are tolerated in siRNA molecules nor provide any examples of such modified siRNA. Kreutzer and Limmer, Canadian Patent Application No. 2,359,180, also
30 describe certain chemical modifications for use in dsRNA constructs in order to counteract activation of double-stranded RNA-dependent protein kinase PKR, specifically 2'-amino or 2'-O-methyl nucleotides, and nucleotides containing a 2'-O or 4'-C methylene bridge. However, Kreutzer and Limmer similarly fail to show to what extent these modifications

are tolerated in siRNA molecules nor do they provide any examples of such modified siRNA.

Parrish et al., 2000, *Molecular Cell*, 6, 1977-1087, tested certain chemical modifications targeting the unc-22 gene in *C. elegans* using long (>25 nt) siRNA transcripts. The authors describe the introduction of thiophosphate residues into these
5 siRNA transcripts by incorporating thiophosphate nucleotide analogs with T7 and T3 RNA polymerase and observed that "RNAs with two [phosphorothioate] modified bases also had substantial decreases in effectiveness as RNAi triggers; [phosphorothioate] modification of more than two residues greatly destabilized the RNAs in vitro and we
10 were not able to assay interference activities." Parrish et al. at 1081. The authors also tested certain modifications at the 2'-position of the nucleotide sugar in the long siRNA transcripts and observed that substituting deoxynucleotides for ribonucleotides "produced a substantial decrease in interference activity", especially in the case of Uridine to Thymidine and/or Cytidine to deoxy-Cytidine substitutions. Parrish et al. In addition, the
15 authors tested certain base modifications, including substituting 4-thiouracil, 5-bromouracil, 5-iodouracil, 3-(aminoallyl)uracil for uracil, and inosine for guanosine in sense and antisense strands of the siRNA, and found that whereas 4-thiouracil and 5-bromouracil were all well tolerated, inosine "produced a substantial decrease in interference activity" when incorporated in either strand. Incorporation of 5-iodouracil and
20 3-(aminoallyl)uracil in the antisense strand resulted in substantial decrease in RNAi activity as well.

Beach et al., WO 01/68836, describes specific methods for attenuating gene expression using endogenously derived dsRNA. Tuschl et al., WO 01/75164, describes a
25 *Drosophila* in vitro RNAi system and the use of specific siRNA molecules for certain functional genomic and certain therapeutic applications; although Tuschl, 2001, *Chem. Biochem.*, 2, 239-245, doubts that RNAi can be used to cure genetic diseases or viral infection due "to the danger of activating interferon response". Li et al., WO 00/44914, describes the use of specific dsRNAs for use in attenuating the expression of certain target genes. Zernicka-Goetz et al., WO 01/36646, describes certain methods for inhibiting the
30 expression of particular genes in mammalian cells using certain dsRNA molecules. Fire et al., WO 99/32619, U.S. Patent No. 6,506,559, the contents of which are hereby incorporated by reference in their entirety, describes particular methods for introducing

certain dsRNA molecules into cells for use in inhibiting gene expression. Plaetinck et al., WO 00/01846, describes certain methods for identifying specific genes responsible for conferring a particular phenotype in a cell using specific dsRNA molecules. Mello et al., WO 01/29058, describes the identification of specific genes involved in dsRNA mediated RNAi. Deschamps Depaillette et al., International PCT Publication No. WO 99/07409, describes specific compositions consisting of particular dsRNA molecules combined with certain anti-viral agents. Driscoll et al., International PCT Publication No. WO 01/49844, describes specific DNA constructs for use in facilitating gene silencing in targeted organisms. Parrish et al., 2000, Molecular Cell, 6, 1977-1087, describes specific chemically modified siRNA constructs targeting the unc-22 gene of *C. elegans*. Tuschl et al., International PCT Publication No. WO 02/44321, describe certain synthetic siRNA constructs.

Methods for Using Nucleic Acid Molecules as Probes and Primers

The isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize, and quantify hybridizing nucleic acids in, and isolate hybridizing nucleic acids from, both genomic and transcript-derived nucleic acid samples. When free in solution, such probes are typically, but not invariably, detectably labeled; bound to a substrate, as in a microarray, such probes are typically, but not invariably unlabeled.

In one embodiment, the isolated nucleic acid molecules of the present invention can be used as probes to detect and characterize gross alterations in the gene of a CSNA, such as deletions, insertions, translocations, and duplications of the CSNA genomic locus through fluorescence *in situ* hybridization (FISH) to chromosome spreads. See, e.g., Andreeff *et al.* (eds.), Introduction to Fluorescence In Situ Hybridization: Principles and Clinical Applications, John Wiley & Sons (1999). The isolated nucleic acid molecules of the present invention can be used as probes to assess smaller genomic alterations using, e.g., Southern blot detection of restriction fragment length polymorphisms. The isolated nucleic acid molecules of the present invention can be used as probes to isolate genomic clones that include a nucleic acid molecule of the present invention, which thereafter can be restriction mapped and sequenced to identify deletions, insertions, translocations, and substitutions (single nucleotide polymorphisms, SNPs) at the sequence level.

Alternatively, detection techniques such as molecular beacons may be used, see Kostrikis *et al. Science* 279:1228-1229 (1998).

The isolated nucleic acid molecules of the present invention can also be used as probes to detect, characterize, and quantify CSNA in, and isolate CSNA from, transcript-derived nucleic acid samples. In one embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by length, and quantify mRNA by Northern blot of total or poly-A⁺- selected RNA samples. In another embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by location, and quantify mRNA by *in situ* hybridization to tissue sections. *See, e.g.,* Schwarchzacher *et al.*, In Situ Hybridization, Springer-Verlag New York (2000). In another preferred embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to measure the representation of clones in a cDNA library or to isolate hybridizing nucleic acid molecules acids from cDNA libraries, permitting sequence level characterization of mRNAs that hybridize to CSNAs, including, without limitations, identification of deletions, insertions, substitutions, truncations, alternatively spliced forms and single nucleotide polymorphisms. In yet another preferred embodiment, the nucleic acid molecules of the instant invention may be used in microarrays.

All of the aforementioned probe techniques are well within the skill in the art, and are described at greater length in standard texts such as Sambrook (2001), *supra*; Ausubel (1999), *supra*; and Walker *et al.* (eds.), The Nucleic Acids Protocols Handbook, Humana Press (2000).

In another embodiment, a nucleic acid molecule of the invention may be used as a probe or primer to identify and/or amplify a second nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of the invention. In this embodiment, it is preferred that the probe or primer be derived from a nucleic acid molecule encoding a CSP. More preferably, the probe or primer is derived from a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 113-259. Also preferred are probes or primers derived from a CSNA. More preferred are probes or primers derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-112.

In general, a probe or primer is at least 10 nucleotides in length, more preferably at least 12, more preferably at least 14 and even more preferably at least 16 or 17 nucleotides

in length. In an even more preferred embodiment, the probe or primer is at least 18 nucleotides in length, even more preferably at least 20 nucleotides and even more preferably at least 22 nucleotides in length. Primers and probes may also be longer in length. For instance, a probe or primer may be 25 nucleotides in length, or may be 30, 40 or 50 nucleotides in length. Methods of performing nucleic acid hybridization using oligonucleotide probes are well known in the art. *See, e.g.*, Sambrook *et al.*, 1989, *supra*, Chapter 11 and pp. 11.31-11.32 and 11.40-11.44, which describes radiolabeling of short probes, and pp. 11.45-11.53, which describe hybridization conditions for oligonucleotide probes, including specific conditions for probe hybridization (pp. 11.50-11.51).

Methods of performing primer-directed amplification are also well known in the art. Methods for performing the polymerase chain reaction (PCR) are compiled, *inter alia*, in McPherson, PCR Basics: From Background to Bench, Springer Verlag (2000); Innis *et al.* (eds.), PCR Applications: Protocols for Functional Genomics, Academic Press (1999); Gelfand *et al.* (eds.), PCR Strategies, Academic Press (1998); Newton *et al.*, PCR, Springer-Verlag New York (1997); Burke (ed.), PCR: Essential Techniques, John Wiley & Son Ltd (1996); White (ed.), PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering, Vol. 67, Humana Press (1996); and McPherson *et al.* (eds.), PCR 2: A Practical Approach, Oxford University Press, Inc. (1995). Methods for performing RT-PCR are collected, *e.g.*, in Siebert *et al.* (eds.), Gene Cloning and Analysis by RT-PCR, Eaton Publishing Company/Bio Techniques Books Division, 1998; and Siebert (ed.), PCR Technique: RT-PCR, Eaton Publishing Company/ BioTechniques Books (1995).

PCR and hybridization methods may be used to identify and/or isolate nucleic acid molecules of the present invention including allelic variants, homologous nucleic acid molecules and fragments. PCR and hybridization methods may also be used to identify, amplify and/or isolate nucleic acid molecules of the present invention that encode homologous proteins, analogs, fusion proteins or muteins of the invention. Nucleic acid primers as described herein can be used to prime amplification of nucleic acid molecules of the invention, using transcript-derived or genomic DNA as the template.

These nucleic acid primers can also be used, for example, to prime single base extension (SBE) for SNP detection (*See, e.g.*, U.S. Pat. No. 6,004,744, the disclosure of which is incorporated herein by reference in its entirety).

Isothermal amplification approaches, such as rolling circle amplification, are also now well-described. *See, e.g.*, Schweitzer *et al.*, *Curr. Opin. Biotechnol.* 12(1): 21-7

(2001); International Patent publications WO 97/19193 and WO 00/15779, and U.S. Patent Nos. 5,854,033 and 5,714,320, the disclosures of which are incorporated herein by reference in their entireties. Rolling circle amplification can be combined with other techniques to facilitate SNP detection. *See, e.g., Lizardi et al., Nature Genet.* 19(3):
5 225-32 (1998).

Nucleic acid molecules of the present invention may be bound to a substrate either covalently or noncovalently. The substrate can be porous or solid, planar or non-planar, unitary or distributed. The bound nucleic acid molecules may be used as hybridization probes, and may be labeled or unlabeled. In a preferred embodiment, the bound nucleic
10 acid molecules are unlabeled.

In one embodiment, the nucleic acid molecule of the present invention is bound to a porous substrate, *e.g.*, a membrane, typically comprising nitrocellulose, nylon, or positively charged derivatized nylon. The nucleic acid molecule of the present invention can be used to detect a hybridizing nucleic acid molecule that is present within a labeled
15 nucleic acid sample, *e.g.*, a sample of transcript-derived nucleic acids. In another embodiment, the nucleic acid molecule is bound to a solid substrate, including, without limitation, glass, amorphous silicon, crystalline silicon or plastics. Examples of plastics include, without limitation, polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene,
20 polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, nitrocellulose, or mixtures thereof. The solid substrate may be any shape, including rectangular, disk-like and spherical. In a preferred embodiment, the solid substrate is a microscope slide or slide-shaped substrate.

The nucleic acid molecule of the present invention can be attached covalently to a
25 surface of the support substrate or applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence by presumed noncovalent interactions, or some combination thereof. The nucleic acid molecule of the present invention can be bound to a substrate to which a plurality of other nucleic acids are concurrently bound, hybridization to each of the plurality of bound nucleic acids being separately detectable. At low density,
30 *e.g.* on a porous membrane, these substrate-bound collections are typically denominated macroarrays; at higher density, typically on a solid support, such as glass, these substrate bound collections of plural nucleic acids are colloquially termed microarrays. As used herein, the term microarray includes arrays of all densities. It is, therefore, another aspect

of the invention to provide microarrays that comprise one or more of the nucleic acid molecules of the present invention.

In yet another embodiment, the invention is directed to single exon probes based on the CSNAs disclosed herein.

5 *Expression Vectors, Host Cells and Recombinant Methods of Producing Polypeptides*

Another aspect of the present invention provides vectors that comprise one or more of the isolated nucleic acid molecules of the present invention, and host cells in which such vectors have been introduced.

10 The vectors can be used, *inter alia*, for propagating the nucleic acid molecules of the present invention in host cells (cloning vectors), for shuttling the nucleic acid molecules of the present invention between host cells derived from disparate organisms (shuttle vectors), for inserting the nucleic acid molecules of the present invention into host cell chromosomes (insertion vectors), for expressing sense or antisense RNA transcripts of
15 the nucleic acid molecules of the present invention *in vitro* or within a host cell, and for expressing polypeptides encoded by the nucleic acid molecules of the present invention, alone or as fusion proteins with heterologous polypeptides (expression vectors). Vectors are by now well known in the art, and are described, *inter alia*, in Jones *et al.* (eds.), Vectors: Cloning Applications: Essential Techniques (Essential Techniques Series), John
20 Wiley & Son Ltd. (1998); Jones *et al.* (eds.), Vectors: Expression Systems: Essential Techniques (Essential Techniques Series), John Wiley & Son Ltd. (1998); Gacesa *et al.*, Vectors: Essential Data, John Wiley & Sons Ltd. (1995); Cid-Arregui (eds.), Viral Vectors: Basic Science and Gene Therapy, Eaton Publishing Co. (2000); Sambrook (2001), *supra*; Ausubel (1999), *supra*. Furthermore, a variety of vectors are available
25 commercially. Use of existing vectors and modifications thereof are well within the skill in the art. Thus, only basic features need be described here.

Nucleic acid sequences may be expressed by operatively linking them to an expression control sequence in an appropriate expression vector and employing that expression vector to transform an appropriate unicellular host. Expression control
30 sequences are sequences that control the transcription, post-transcriptional events and translation of nucleic acid sequences. Such operative linking of a nucleic acid sequence of this invention to an expression control sequence, of course, includes, if not already part of

the nucleic acid sequence, the provision of a translation initiation codon, ATG or GTG, in the correct reading frame upstream of the nucleic acid sequence.

A wide variety of host/expression vector combinations may be employed in expressing the nucleic acid sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic nucleic acid sequences.

In one embodiment, prokaryotic cells may be used with an appropriate vector. Prokaryotic host cells are often used for cloning and expression. In a preferred embodiment, prokaryotic host cells include *E. coli*, *Pseudomonas*, *Bacillus* and *Streptomyces*. In a preferred embodiment, bacterial host cells are used to express the nucleic acid molecules of the instant invention. Useful expression vectors for bacterial hosts include bacterial plasmids, such as those from *E. coli*, *Bacillus* or *Streptomyces*, including pBluescript, pGEX-2T, pUC vectors, col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, such as RP4, phage DNAs, *e.g.*, the numerous derivatives of phage lambda, *e.g.*, NM989, λ GT10 and λ GT11, and other phages, *e.g.*, M13 and filamentous single-stranded phage DNA. Where *E. coli* is used as host, selectable markers are, analogously, chosen for selectivity in gram negative bacteria: *e.g.*, typical markers confer resistance to antibiotics, such as ampicillin, tetracycline, chloramphenicol, kanamycin, streptomycin and zeocin; auxotrophic markers can also be used.

In other embodiments, eukaryotic host cells, such as yeast, insect, mammalian or plant cells, may be used. Yeast cells, typically *S. cerevisiae*, are useful for eukaryotic genetic studies, due to the ease of targeting genetic changes by homologous recombination and the ability to easily complement genetic defects using recombinantly expressed proteins. Yeast cells are useful for identifying interacting protein components, *e.g.* through use of a two-hybrid system. In a preferred embodiment, yeast cells are useful for protein expression. Vectors of the present invention for use in yeast will typically, but not invariably, contain an origin of replication suitable for use in yeast and a selectable marker that is functional in yeast. Yeast vectors include Yeast Integrating plasmids (*e.g.*, YIp5) and Yeast Replicating plasmids (the YRp and YEplac series plasmids), Yeast Centromere plasmids (the YCp series plasmids), Yeast Artificial Chromosomes (YACs) which are based on yeast linear plasmids, denoted YLp, pGPD-2, 2 μ plasmids and derivatives thereof, and improved shuttle vectors such as those described in Gietz *et al.*, *Gene*, 74:

527-34 (1988) (YIplac, YEplac and YCplac). Selectable markers in yeast vectors include a variety of auxotrophic markers, the most common of which are (in *Saccharomyces cerevisiae*) URA3, HIS3, LEU2, TRP1 and LYS2, which complement specific auxotrophic mutations, such as *ura3-52*, *his3-D1*, *leu2-D1*, *trp1-D1* and *lys2-201*.

- 5 Insect cells may be chosen for high efficiency protein expression. Where the host cells are from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and expresSFT™ cells (Protein Sciences Corp., Meriden, CT, USA), the vector replicative strategy is typically based upon the baculovirus life cycle. Typically, baculovirus transfer vectors are used to replace the wild-type AcMNPV polyhedrin gene with a heterologous gene of interest.
- 10 Sequences that flank the polyhedrin gene in the wild-type genome are positioned 5' and 3' of the expression cassette on the transfer vectors. Following co-transfection with AcMNPV DNA, a homologous recombination event occurs between these sequences resulting in a recombinant virus carrying the gene of interest and the polyhedrin or p10 promoter. Selection can be based upon visual screening for lacZ fusion activity.
- 15 The host cells may also be mammalian cells, which are particularly useful for expression of proteins intended as pharmaceutical agents, and for screening of potential agonists and antagonists of a protein or a physiological pathway. Mammalian vectors intended for autonomous extrachromosomal replication will typically include a viral origin, such as the SV40 origin (for replication in cell lines expressing the large T-antigen,
- 20 such as COS1 and COS7 cells), the papillomavirus origin, or the EBV origin for long term episomal replication (for use, e.g., in 293-EBNA cells, which constitutively express the EBV EBNA-1 gene product and adenovirus E1A). Vectors intended for integration, and thus replication as part of the mammalian chromosome, can, but need not, include an origin of replication functional in mammalian cells, such as the SV40 origin. Vectors
- 25 based upon viruses, such as adenovirus, adeno-associated virus, vaccinia virus, and various mammalian retroviruses, will typically replicate according to the viral replicative strategy. Selectable markers for use in mammalian cells include, but are not limited to, resistance to neomycin (G418), blasticidin, hygromycin and zeocin, and selection based upon the purine salvage pathway using HAT medium.
- 30 Expression in mammalian cells can be achieved using a variety of plasmids, including pSV2, pBC12BI, and p91023, as well as lytic virus vectors (e.g., vaccinia virus, adeno virus, and baculovirus), episomal virus vectors (e.g., bovine papillomavirus), and

retroviral vectors (*e.g.*, murine retroviruses). Useful vectors for insect cells include baculoviral vectors and pVL 941.

Plant cells can also be used for expression, with the vector replicon typically derived from a plant virus (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, 5 TMV) and selectable markers chosen for suitability in plants.

It is known that codon usage of different host cells may be different. For example, a plant cell and a human cell may exhibit a difference in codon preference for encoding a particular amino acid. As a result, human mRNA may not be efficiently translated in a plant, bacteria or insect host cell. Therefore, another embodiment of this invention is 10 directed to codon optimization. The codons of the nucleic acid molecules of the invention may be modified to resemble, as much as possible, genes naturally contained within the host cell without altering the amino acid sequence encoded by the nucleic acid molecule.

Any of a wide variety of expression control sequences may be used in these vectors to express the nucleic acid molecules of this invention. Such useful expression 15 control sequences include the expression control sequences associated with structural genes of the foregoing expression vectors. Expression control sequences that control transcription include, *e.g.*, promoters, enhancers and transcription termination sites. Expression control sequences in eukaryotic cells that control post-transcriptional events include splice donor and acceptor sites and sequences that modify the half-life of the 20 transcribed RNA, *e.g.*, sequences that direct poly(A) addition or binding sites for RNA-binding proteins. Expression control sequences that control translation include ribosome binding sites, sequences which direct targeted expression of the polypeptide to or within particular cellular compartments, and sequences in the 5' and 3' untranslated regions that modify the rate or efficiency of translation.

25 Examples of useful expression control sequences for a prokaryote, *e.g.*, *E. coli*, will include a promoter, often a phage promoter, such as phage lambda pL promoter, the trc promoter, a hybrid derived from the trp and lac promoters, the bacteriophage T7 promoter (in *E. coli* cells engineered to express the T7 polymerase), the TAC or TRC system, the major operator and promoter regions of phage lambda, the control regions of 30 fd coat protein, and the araBAD operon. Prokaryotic expression vectors may further include transcription terminators, such as the aspA terminator, and elements that facilitate translation, such as a consensus ribosome binding site and translation termination codon, Schomer *et al.*, *Proc. Natl. Acad. Sci. USA* 83: 8506-8510 (1986).

Expression control sequences for yeast cells, typically *S. cerevisiae*, will include a yeast promoter, such as the CYC1 promoter, the GAL1 promoter, the GAL10 promoter, ADH1 promoter, the promoters of the yeast α -mating system, or the GPD promoter, and will typically have elements that facilitate transcription termination, such as the
5 transcription termination signals from the CYC1 or ADH1 gene.

Expression vectors useful for expressing proteins in mammalian cells will include a promoter active in mammalian cells. These promoters include, but are not limited to, those derived from mammalian viruses, such as the enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), the enhancer-promoter
10 sequences from the Rous sarcoma virus long terminal repeat (RSV LTR), the enhancer-promoter from SV40 and the early and late promoters of adenovirus. Other expression control sequences include the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase. Other expression control sequences include those from the gene comprising the CSNA of interest. Often, expression is enhanced by
15 incorporation of polyadenylation sites, such as the late SV40 polyadenylation site and the polyadenylation signal and transcription termination sequences from the bovine growth hormone (BGH) gene, and ribosome binding sites. Furthermore, vectors can include introns, such as intron II of rabbit β -globin gene and the SV40 splice elements.

Preferred nucleic acid vectors also include a selectable or amplifiable marker gene
20 and means for amplifying the copy number of the gene of interest. Such marker genes are well known in the art. Nucleic acid vectors may also comprise stabilizing sequences (*e.g.*, ori- or ARS-like sequences and telomere-like sequences), or may alternatively be designed to favor directed or non-directed integration into the host cell genome. In a preferred embodiment, nucleic acid sequences of this invention are inserted in frame into an
25 expression vector that allows a high level expression of an RNA which encodes a protein comprising the encoded nucleic acid sequence of interest. Nucleic acid cloning and sequencing methods are well known to those of skill in the art and are described in an assortment of laboratory manuals, including Sambrook (1989), *supra*, Sambrook (2000), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Product information from
30 manufacturers of biological, chemical and immunological reagents also provide useful information.

Expression vectors may be either constitutive or inducible. Inducible vectors include either naturally inducible promoters, such as the *trc* promoter, which is regulated

by the lac operon, and the pL promoter, which is regulated by tryptophan, the MMTV-LTR promoter, which is inducible by dexamethasone, or can contain synthetic promoters and/or additional elements that confer inducible control on adjacent promoters. Examples of inducible synthetic promoters are the hybrid Plac/ara-1 promoter and the

5 PLtetO-1 promoter. The PLtetO-1 promoter takes advantage of the high expression levels from the PL promoter of phage lambda, but replaces the lambda repressor sites with two copies of operator 2 of the Tn10 tetracycline resistance operon, causing this promoter to be tightly repressed by the Tet repressor protein and induced in response to tetracycline (Tc) and Tc derivatives such as anhydrotetracycline. Vectors may also be inducible

10 because they contain hormone response elements, such as the glucocorticoid response element (GRE) and the estrogen response element (ERE), which can confer hormone inducibility where vectors are used for expression in cells having the respective hormone receptors. To reduce background levels of expression, elements responsive to ecdysone, an insect hormone, can be used instead, with coexpression of the ecdysone receptor.

15 In one embodiment of the invention, expression vectors can be designed to fuse the expressed polypeptide to small protein tags that facilitate purification and/or visualization. Such tags include a polyhistidine tag that facilitates purification of the fusion protein by immobilized metal affinity chromatography, for example using NiNTA resin (Qiagen Inc., Valencia, CA, USA) or TALON™ resin (cobalt immobilized affinity chromatography

20 medium, Clontech Labs, Palo Alto, CA, USA). The fusion protein can include a chitin-binding tag and self-excising intein, permitting chitin-based purification with self-removal of the fused tag (IMPACT™ system, New England Biolabs, Inc., Beverly, MA, USA). Alternatively, the fusion protein can include a calmodulin-binding peptide tag, permitting purification by calmodulin affinity resin (Stratagene, La Jolla, CA, USA), or a specifically

25 excisable fragment of the biotin carboxylase carrier protein, permitting purification of *in vivo* biotinylated protein using an avidin resin and subsequent tag removal (Promega, Madison, WI, USA). As another useful alternative, the polypeptides of the present invention can be expressed as a fusion to glutathione-S-transferase, the affinity and specificity of binding to glutathione permitting purification using glutathione affinity

30 resins, such as Glutathione-Superflow Resin (Clontech Laboratories, Palo Alto, CA, USA), with subsequent elution with free glutathione. Other tags include, for example, the Xpress epitope, detectable by anti-Xpress antibody (Invitrogen, Carlsbad, CA, USA), a myc tag, detectable by anti-myc tag antibody, the V5 epitope, detectable by anti-V5

antibody (Invitrogen, Carlsbad, CA, USA), FLAG® epitope, detectable by anti-FLAG® antibody (Stratagene, La Jolla, CA, USA), and the HA epitope, detectable by anti-HA antibody.

For secretion of expressed polypeptides, vectors can include appropriate sequences
5 that encode secretion signals, such as leader peptides. For example, the pSecTag2 vectors (Invitrogen, Carlsbad, CA, USA) are 5.2 kb mammalian expression vectors that carry the secretion signal from the V-J2-C region of the mouse Ig kappa-chain for efficient secretion of recombinant proteins from a variety of mammalian cell lines.

Expression vectors can also be designed to fuse proteins encoded by the
10 heterologous nucleic acid insert to polypeptides that are larger than purification and/or identification tags. Useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as those that have a green fluorescent protein (GFP)-like chromophore, fusions to the IgG Fc region, and fusions for use in two hybrid systems.

15 Vectors for phage display fuse the encoded polypeptide to, *e.g.*, the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13. See Barbas *et al.*, Phage Display: A Laboratory Manual, Cold Spring Harbor Laboratory Press (2001); Kay *et al.* (eds.), Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press, Inc., (1996); Abelson *et al.* (eds.), Combinatorial
20 Chemistry (Methods in Enzymology, Vol. 267) Academic Press (1996). Vectors for yeast display, *e.g.* the pYD1 yeast display vector (Invitrogen, Carlsbad, CA, USA), use the α -agglutinin yeast adhesion receptor to display recombinant protein on the surface of *S. cerevisiae*. Vectors for mammalian display, *e.g.*, the pDisplay™ vector (Invitrogen, Carlsbad, CA, USA), target recombinant proteins using an N-terminal cell surface
25 targeting signal and a C-terminal transmembrane anchoring domain of platelet derived growth factor receptor.

A wide variety of vectors now exist that fuse proteins encoded by heterologous nucleic acids to the chromophore of the substrate-independent, intrinsically fluorescent green fluorescent protein from *Aequorea victoria* ("GFP") and its variants. The GFP-like
30 chromophore can be selected from GFP-like chromophores found in naturally occurring proteins, such as *A. victoria* GFP (GenBank accession number AAA27721), *Renilla reniformis* GFP, FP583 (GenBank accession no. AF168419) (DsRed), FP593 (AF272711), FP483 (AF168420), FP484 (AF168424), FP595 (AF246709), FP486 (AF168421), FP538

(AF168423), and FP506 (AF168422), and need include only so much of the native protein as is needed to retain the chromophore's intrinsic fluorescence. Methods for determining the minimal domain required for fluorescence are known in the art. *See Li et al., J. Biol. Chem.* 272: 28545-28549 (1997). Alternatively, the GFP-like chromophore can be

5 selected from GFP-like chromophores modified from those found in nature. The methods for engineering such modified GFP-like chromophores and testing them for fluorescence activity, both alone and as part of protein fusions, are well known in the art. *See Heim et al., Curr. Biol.* 6: 178-182 (1996) and Palm *et al., Methods Enzymol.* 302: 378-394 (1999). A variety of such modified chromophores are now commercially available and can readily

10 be used in the fusion proteins of the present invention. These include EGFP ("enhanced GFP"), EBFP ("enhanced blue fluorescent protein"), BFP2, EYFP ("enhanced yellow fluorescent protein"), ECFP ("enhanced cyan fluorescent protein") or Citrine. EGFP (*see, e.g., Cormack et al., Gene* 173: 33-38 (1996); U.S. Patent Nos. 6,090,919 and 5,804,387, the disclosures of which are incorporated herein by reference in their entireties) is found

15 on a variety of vectors, both plasmid and viral, which are available commercially (Clontech Labs, Palo Alto, CA, USA); EBFP is optimized for expression in mammalian cells whereas BFP2, which retains the original jellyfish codons, can be expressed in bacteria (*see, e.g., Heim et al., Curr. Biol.* 6: 178-182 (1996) and Cormack *et al., Gene* 173: 33-38 (1996)). Vectors containing these blue-shifted variants are available from

20 Clontech Labs (Palo Alto, CA, USA). Vectors containing EYFP, ECFP (*see, e.g., Heim et al., Curr. Biol.* 6: 178-182 (1996); Miyawaki *et al., Nature* 388: 882-887 (1997)) and Citrine (*see, e.g., Heikal et al., Proc. Natl. Acad. Sci. USA* 97: 11996-12001 (2000)) are also available from Clontech Labs. The GFP-like chromophore can also be drawn from other modified GFPs, including those described in U.S. Patent Nos. 6,124,128; 6,096,865;

25 6,090,919; 6,066,476; 6,054,321; 6,027,881; 5,968,750; 5,874,304; 5,804,387; 5,777,079; 5,741,668; and 5,625,048, the disclosures of which are incorporated herein by reference in their entireties. *See also Conn (ed.), Green Fluorescent Protein* (Methods in Enzymology, Vol. 302), Academic Press, Inc. (1999); Yang, *et al., J Biol Chem*, 273: 8212-6 (1998); Bevis *et al., Nature Biotechnology*, 20:83-7 (2002). The GFP-like chromophore of each

30 of these GFP variants can usefully be included in the fusion proteins of the present invention.

Fusions to the IgG Fc region increase serum half-life of protein pharmaceutical products through interaction with the FcRn receptor (also denominated the FcRp receptor

and the Brambell receptor, FcRb), further described in International Patent Application Nos. WO 97/43316, WO 97/34631, WO 96/32478, and WO 96/18412, the disclosures of which are incorporated herein by reference in their entireties.

For long-term, high-yield recombinant production of the polypeptides of the present invention, stable expression is preferred. Stable expression is readily achieved by integration into the host cell genome of vectors having selectable markers, followed by selection of these integrants. Vectors such as pUB6/V5-His A, B, and C (Invitrogen, Carlsbad, CA, USA) are designed for high-level stable expression of heterologous proteins in a wide range of mammalian tissue types and cell lines. pUB6/V5-His uses the promoter/enhancer sequence from the human ubiquitin C gene to drive expression of recombinant proteins: expression levels in 293, CHO, and NIH3T3 cells are comparable to levels from the CMV and human EF-1a promoters. The bsd gene permits rapid selection of stably transfected mammalian cells with the potent antibiotic blasticidin.

Replication incompetent retroviral vectors, typically derived from Moloney murine leukemia virus, also are useful for creating stable transfectants having integrated provirus. The highly efficient transduction machinery of retroviruses, coupled with the availability of a variety of packaging cell lines such as RetroPack™ PT 67, EcoPack2™-293, AmphoPack-293, and GP2-293 cell lines (all available from Clontech Laboratories, Palo Alto, CA, USA) allow a wide host range to be infected with high efficiency; varying the multiplicity of infection readily adjusts the copy number of the integrated provirus.

Of course, not all vectors and expression control sequences will function equally well to express the nucleic acid molecules of this invention. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences and hosts without undue experimentation and without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must be replicated in it. The vector's copy number, the ability to control that copy number, the ability to control integration, if any, and the expression of any other proteins encoded by the vector, such as an antibiotic or other selection marker, should also be considered. The present invention further includes host cells comprising the vectors of the present invention, either present episomally within the cell or integrated, in whole or in part, into the host cell chromosome. Among other considerations, some of which are described above, a host cell strain may be chosen for its ability to process the expressed polypeptide in the desired fashion. Such

post-translational modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation, and it is an aspect of the present invention to provide CSPs with such post-translational modifications.

5 In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the sequence, its controllability, and its compatibility with the nucleic acid molecules of this invention, particularly with regard to potential secondary structures. Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the
10 product coded for by the nucleic acid sequences of this invention, their secretion characteristics, their ability to fold the polypeptide correctly, their fermentation or culture requirements, and the ease of purification from them of the products coded for by the nucleic acid molecules of this invention.

The recombinant nucleic acid molecules and more particularly, the expression
15 vectors of this invention may be used to express the polypeptides of this invention as recombinant polypeptides in a heterologous host cell. The polypeptides of this invention may be full-length or less than full-length polypeptide fragments recombinantly expressed from the nucleic acid molecules according to this invention. Such polypeptides include analogs, derivatives and muteins that may or may not have biological activity.

20 Vectors of the present invention will also often include elements that permit *in vitro* transcription of RNA from the inserted heterologous nucleic acid. Such vectors typically include a phage promoter, such as that from T7, T3, or SP6, flanking the nucleic acid insert. Often two different such promoters flank the inserted nucleic acid, permitting separate *in vitro* production of both sense and antisense strands.

25 Transformation and other methods of introducing nucleic acids into a host cell (e.g., conjugation, protoplast transformation or fusion, transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion) can be accomplished by a variety of methods which are well known in the art (See, for instance, Ausubel, *supra*, and Sambrook *et al.*, *supra*).

30 Bacterial, yeast, plant or mammalian cells are transformed or transfected with an expression vector, such as a plasmid, a cosmid, or the like, wherein the expression vector comprises the nucleic acid of interest. Alternatively, the cells may be infected by a viral expression vector comprising the nucleic acid of interest. Depending upon the host cell,

vector, and method of transformation used, transient or stable expression of the polypeptide will be constitutive or inducible. One having ordinary skill in the art will be able to decide whether to express a polypeptide transiently or stably, and whether to express the protein constitutively or inducibly.

- 5 A wide variety of unicellular host cells are useful in expressing the DNA sequences of this invention. These hosts may include well known eukaryotic and prokaryotic hosts, such as strains of, fungi, yeast, insect cells such as *Spodoptera frugiperda* (SF9), animal cells such as CHO, as well as plant cells in tissue culture. Representative examples of appropriate host cells include, but are not limited to, bacterial
- 10 cells, such as *E. coli*, *Caulobacter crescentus*, *Streptomyces* species, and *Salmonella typhimurium*; yeast cells, such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Pichia methanolica*; insect cell lines, such as those from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and expresSF™ cells (Protein Sciences Corp., Meriden, CT, USA), *Drosophila* S2 cells, and *Trichoplusia ni* High Five® Cells
- 15 (Invitrogen, Carlsbad, CA, USA); and mammalian cells. Typical mammalian cells include BHK cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, COS1 cells, COS7 cells, Chinese hamster ovary (CHO) cells, 3T3 cells, NIH 3T3 cells, 293 cells, HEPG2 cells, HeLa cells, L cells, MDCK cells, HEK293 cells, WI38 cells, murine ES cell lines (e.g., from strains 129/SV, C57/BL6, DBA-1, 129/SVJ), K562 cells, Jurkat cells, and
- 20 BW5147 cells. Other mammalian cell lines are well known and readily available from the American Type Culture Collection (ATCC) (Manassas, VA, USA) and the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA). Cells or cell lines derived from colon are particularly preferred because they may provide a more native post-translational
- 25 processing. Particularly preferred are human colon cells.

- Particular details of the transfection, expression and purification of recombinant proteins are well documented and are understood by those of skill in the art. Further details on the various technical aspects of each of the steps used in recombinant production of foreign genes in bacterial cell expression systems can be found in a number
- 30 of texts and laboratory manuals in the art. See, e.g., Ausubel (1992), *supra*, Ausubel (1999), *supra*, Sambrook (1989), *supra*, and Sambrook (2001), *supra*.

Methods for introducing the vectors and nucleic acid molecules of the present invention into the host cells are well known in the art; the choice of technique will depend primarily upon the specific vector to be introduced and the host cell chosen.

Nucleic acid molecules and vectors may be introduced into prokaryotes, such as *E. coli*, in a number of ways. For instance, phage lambda vectors will typically be packaged using a packaging extract (e.g., Gigapack® packaging extract, Stratagene, La Jolla, CA, USA), and the packaged virus used to infect *E. coli*.

Plasmid vectors will typically be introduced into chemically competent or electrocompetent bacterial cells. *E. coli* cells can be rendered chemically competent by treatment, e.g., with CaCl_2 , or a solution of Mg^{2+} , Mn^{2+} , Ca^{2+} , Rb^+ or K^+ , dimethyl sulfoxide, dithiothreitol, and hexamine cobalt (III), Hanahan, *J. Mol. Biol.* 166(4):557-80 (1983), and vectors introduced by heat shock. A wide variety of chemically competent strains are also available commercially (e.g., Epicurian Coli® XL10-Gold® Ultracompetent Cells (Stratagene, La Jolla, CA, USA); DH5α competent cells (Clontech Laboratories, Palo Alto, CA, USA); and TOP10 Chemically Competent *E. coli* Kit (Invitrogen, Carlsbad, CA, USA)). Bacterial cells can be rendered electrocompetent to take up exogenous DNA by electroporation by various pre-pulse treatments; vectors are introduced by electroporation followed by subsequent outgrowth in selected media. An extensive series of protocols is provided by BioRad (Richmond, CA, USA).

Vectors can be introduced into yeast cells by spheroplasting, treatment with lithium salts, electroporation, or protoplast fusion. Spheroplasts are prepared by the action of hydrolytic enzymes such as a snail-gut extract, usually denoted Glusulase or Zymolyase, or an enzyme from *Arthrobacter luteus* to remove portions of the cell wall in the presence of osmotic stabilizers, typically 1 M sorbitol. DNA is added to the spheroplasts, and the mixture is co-precipitated with a solution of polyethylene glycol (PEG) and Ca^{2+} . Subsequently, the cells are resuspended in a solution of sorbitol, mixed with molten agar and then layered on the surface of a selective plate containing sorbitol.

For lithium-mediated transformation, yeast cells are treated with lithium acetate to permeabilize the cell wall, DNA is added and the cells are co-precipitated with PEG. The cells are exposed to a brief heat shock, washed free of PEG and lithium acetate, and subsequently spread on plates containing ordinary selective medium. Increased frequencies of transformation are obtained by using specially-prepared single-stranded

carrier DNA and certain organic solvents. Schiestl *et al.*, *Curr. Genet.* 16(5-6): 339-46 (1989).

For electroporation, freshly-grown yeast cultures are typically washed, suspended in an osmotic protectant, such as sorbitol, mixed with DNA, and the cell suspension pulsed in an electroporation device. Subsequently, the cells are spread on the surface of plates containing selective media. Becker *et al.*, *Methods Enzymol.* 194: 182-187 (1991). The efficiency of transformation by electroporation can be increased over 100-fold by using PEG, single-stranded carrier DNA and cells that are in late log-phase of growth. Larger constructs, such as YACs, can be introduced by protoplast fusion.

Mammalian and insect cells can be directly infected by packaged viral vectors, or transfected by chemical or electrical means. For chemical transfection, DNA can be coprecipitated with CaPO₄ or introduced using liposomal and nonliposomal lipid-based agents. Commercial kits are available for CaPO₄ transfection (CalPhos™ Mammalian Transfection Kit, Clontech Laboratories, Palo Alto, CA, USA), and lipid-mediated transfection can be practiced using commercial reagents, such as LIPOFECTAMINE™ 2000, LIPOFECTAMINE™ Reagent, CELLFECTIN® Reagent, and LIPOFECTIN® Reagent (Invitrogen, Carlsbad, CA, USA), DOTAP Liposomal Transfection Reagent, FuGENE 6, X-tremeGENE Q2, DOSPER, (Roche Molecular Biochemicals, Indianapolis, IN USA), Effectene™, PolyFect®, Superfect® (Qiagen, Inc., Valencia, CA, USA). Protocols for electroporating mammalian cells can be found in, for example, ; Norton *et al.* (eds.), Gene Transfer Methods: Introducing DNA into Living Cells and Organisms, BioTechniques Books, Eaton Publishing Co. (2000). Other transfection techniques include transfection by particle bombardment and microinjection. See, e.g., Cheng *et al.*, *Proc. Natl. Acad. Sci. USA* 90(10): 4455-9 (1993); Yang *et al.*, *Proc. Natl. Acad. Sci. USA* 87(24): 9568-72 (1990).

Production of the recombinantly produced proteins of the present invention can optionally be followed by purification.

Purification of recombinantly expressed proteins is now well within the skill in the art and thus need not be detailed here. See, e.g., Thorner *et al.* (eds.), Applications of Chimeric Genes and Hybrid Proteins, Part A: Gene Expression and Protein Purification (Methods in Enzymology, Vol. 326), Academic Press (2000); Harbin (ed.), Cloning, Gene Expression and Protein Purification : Experimental Procedures and Process Rationale, Oxford Univ. Press (2001); Marshak *et al.*, Strategies for Protein Purification and

Characterization: A Laboratory Course Manual, Cold Spring Harbor Laboratory Press (1996); and Roe (ed.), Protein Purification Applications, Oxford University Press (2001).

Briefly, however, if purification tags have been fused through use of an expression vector that appends such tags, purification can be effected, at least in part, by means
5 appropriate to the tag, such as use of immobilized metal affinity chromatography for polyhistidine tags. Other techniques common in the art include ammonium sulfate fractionation, immunoprecipitation, fast protein liquid chromatography (FPLC), high performance liquid chromatography (HPLC), and preparative gel electrophoresis.

10 Polypeptides, including Fragments Muteins, Homologous Proteins, Allelic Variants, Analogs and Derivatives

Another aspect of the invention relates to polypeptides encoded by the nucleic acid molecules described herein. In a preferred embodiment, the polypeptide is a colon specific polypeptide (CSP). In an even more preferred embodiment, the polypeptide comprises an amino acid sequence of SEQ ID NO:113-259 or is derived from a
15 polypeptide having the amino acid sequence of SEQ ID NO: 113-259. A polypeptide as defined herein may be produced recombinantly, as discussed *supra*, may be isolated from a cell that naturally expresses the protein, or may be chemically synthesized following the teachings of the specification and using methods well known to those having ordinary skill in the art.

20 Polypeptides of the present invention may also comprise a part or fragment of a CSP. In a preferred embodiment, the fragment is derived from a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 113-259. Polypeptides of the present invention comprising a part or fragment of an entire CSP may or may not be CSPs. For example, a full-length polypeptide may be colon-specific, while
25 a fragment thereof may be found in other tissues as well as in colon. A polypeptide that is not a CSP, whether it is a fragment, analog, mutein, homologous protein or derivative, is nevertheless useful, especially for immunizing animals to prepare anti-CSP antibodies. In a preferred embodiment, the part or fragment is a CSP. Methods of determining whether a polypeptide of the present invention is a CSP are described *infra*.

30 Polypeptides of the present invention comprising fragments of at least 6 contiguous amino acids are also useful in mapping B cell and T cell epitopes of the reference protein. *See, e.g., Geysen et al., Proc. Natl. Acad. Sci. USA 81: 3998-4002*

(1984) and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. Because the fragment need not itself be immunogenic, part of an immunodominant epitope, nor even recognized by native antibody, to be useful in such epitope mapping, all fragments of at least 6 amino acids of a polypeptide of the present invention have utility in such a study.

Polypeptides of the present invention comprising fragments of at least 8 contiguous amino acids, often at least 15 contiguous amino acids, are useful as immunogens for raising antibodies that recognize polypeptides of the present invention. *See, e.g.,* Lerner, *Nature* 299: 592-596 (1982); Shinnick *et al.*, *Annu. Rev. Microbiol.* 37: 425-46 (1983); Sutcliffe *et al.*, *Science* 219: 660-6 (1983). As further described in the above-cited references, virtually all 8-mers, conjugated to a carrier, such as a protein, prove immunogenic and are capable of eliciting antibody for the conjugated peptide; accordingly, all fragments of at least 8 amino acids of the polypeptides of the present invention have utility as immunogens.

Polypeptides comprising fragments of at least 8, 9, 10 or 12 contiguous amino acids are also useful as competitive inhibitors of binding of the entire polypeptide, or a portion thereof, to antibodies (as in epitope mapping), and to natural binding partners, such as subunits in a multimeric complex or to receptors or ligands of the subject protein; this competitive inhibition permits identification and separation of molecules that bind specifically to the polypeptide of interest. *See* U.S. Patent Nos. 5,539,084 and 5,783,674, incorporated herein by reference in their entireties.

The polypeptide of the present invention thus preferably is at least 6 amino acids in length, typically at least 8, 9, 10 or 12 amino acids in length, and often at least 15 amino acids in length. Often, the polypeptide of the present invention is at least 20 amino acids in length, even 25 amino acids, 30 amino acids, 35 amino acids, or 50 amino acids or more in length. Of course, larger polypeptides having at least 75 amino acids, 100 amino acids, or even 150 amino acids are also useful, and at times preferred.

One having ordinary skill in the art can produce fragments by truncating the nucleic acid molecule, *e.g.,* a CSNA, encoding the polypeptide and then expressing it recombinantly. Alternatively, one can produce a fragment by chemically synthesizing a portion of the full-length polypeptide. One may also produce a fragment by enzymatically cleaving either a recombinant polypeptide or an isolated naturally occurring polypeptide. Methods of producing polypeptide fragments are well known in the art. *See, e.g.,*

Sambrook (1989), *supra*; Sambrook (2001), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. In one embodiment, a polypeptide comprising only a fragment, preferably a fragment of a CSP, may be produced by chemical or enzymatic cleavage of a CSP polypeptide. In a preferred embodiment, a polypeptide fragment is produced by
5 expressing a nucleic acid molecule of the present invention encoding a fragment, preferably of a CSP, in a host cell.

Polypeptides of the present invention are also inclusive of mutants, fusion proteins, homologous proteins and allelic variants.

A mutant protein, or mutein, may have the same or different properties compared
10 to a naturally occurring polypeptide and comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of a native polypeptide. Small deletions and insertions can often be found that do not alter the function of a protein. Muteins may or may not be colon-specific. Preferably, the mutein is colon-specific. More preferably the mutein is a polypeptide that comprises at
15 least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of SEQ ID NO: 113-259. Accordingly, in a preferred embodiment, the mutein is one that exhibits at least 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more preferably at least 80% sequence identity to a CSP comprising an amino acid sequence of
20 SEQ ID NO: 113-259. In a yet more preferred embodiment, the mutein exhibits at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97%, 98%, 99% or 99.5% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 113-259.

A mutein may be produced by isolation from a naturally occurring mutant cell,
25 tissue or organism. A mutein may be produced by isolation from a cell, tissue or organism that has been experimentally mutagenized. Alternatively, a mutein may be produced by chemical manipulation of a polypeptide, such as by altering the amino acid residue to another amino acid residue using synthetic or semi-synthetic chemical techniques. In a preferred embodiment, a mutein is produced from a host cell comprising a mutated nucleic
30 acid molecule compared to the naturally occurring nucleic acid molecule. For instance, one may produce a mutein of a polypeptide by introducing one or more mutations into a nucleic acid molecule of the invention and then expressing it recombinantly. These mutations may be targeted, in which particular encoded amino acids are altered, or may be

untargeted, in which random encoded amino acids within the polypeptide are altered. Muteins with random amino acid alterations can be screened for a particular biological activity or property, particularly whether the polypeptide is colon-specific, as described below. Multiple random mutations can be introduced into the gene by methods well known to the art, *e.g.*, by error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis and site-specific mutagenesis. Methods of producing muteins with targeted or random amino acid alterations are well known in the art. *See, e.g.*, Sambrook (1989), *supra*; Sambrook (2001), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), as well as U.S. Patent No. 5,223,408, which is herein incorporated by reference in its entirety.

The invention also contemplates polypeptides that are homologous to a polypeptide of the invention. In a preferred embodiment, the polypeptide is homologous to a CSP. In an even more preferred embodiment, the polypeptide is homologous to a CSP selected from the group having an amino acid sequence of SEQ ID NO: 113-259. By homologous polypeptide it is meant one that exhibits significant sequence identity to a CSP, preferably a CSP having an amino acid sequence of SEQ ID NO: 113-259. By significant sequence identity it is meant that the homologous polypeptide exhibits at least 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more preferably at least 80% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 113-259. More preferred are homologous polypeptides exhibiting at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97% or 98% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 113-259. Most preferably, the homologous polypeptide exhibits at least 99%, more preferably 99.5%, even more preferably 99.6%, 99.7%, 99.8% or 99.9% sequence identity to a CSP comprising an amino acid sequence of SEQ ID NO: 113-259. In a preferred embodiment, the amino acid substitutions of the homologous polypeptide are conservative amino acid substitutions as discussed *supra*.

Homologous polypeptides of the present invention also comprise polypeptide encoded by a nucleic acid molecule that selectively hybridizes to a CSNA or an antisense sequence thereof. In this embodiment, it is preferred that the homologous polypeptide be encoded by a nucleic acid molecule that hybridizes to a CSNA under low stringency,

moderate stringency or high stringency conditions, as defined herein. More preferred is a homologous polypeptide encoded by a nucleic acid sequence which hybridizes to a CSNA selected from the group consisting of SEQ ID NO: 1-112 or a homologous polypeptide encoded by a nucleic acid molecule that hybridizes to a nucleic acid molecule that encodes a CSP, preferably a CSP of SEQ ID NO: 113-259 under low stringency, moderate stringency or high stringency conditions, as defined herein.

Homologous polypeptides of the present invention may be naturally occurring and derived from another species, especially one derived from another primate, such as chimpanzee, gorilla, rhesus macaque, or baboon, wherein the homologous polypeptide comprises an amino acid sequence that exhibits significant sequence identity to that of SEQ ID NO: 113-259. The homologous polypeptide may also be a naturally occurring polypeptide from a human, when the CSP is a member of a family of polypeptides. The homologous polypeptide may also be a naturally occurring polypeptide derived from a non-primate, mammalian species, including without limitation, domesticated species, *e.g.*, dog, cat, mouse, rat, rabbit, guinea pig, hamster, cow, horse, goat or pig. The homologous polypeptide may also be a naturally occurring polypeptide derived from a non-mammalian species, such as birds or reptiles. The naturally occurring homologous protein may be isolated directly from humans or other species. Alternatively, the nucleic acid molecule encoding the naturally occurring homologous polypeptide may be isolated and used to express the homologous polypeptide recombinantly. The homologous polypeptide may also be one that is experimentally produced by random mutation of a nucleic acid molecule and subsequent expression of the nucleic acid molecule. Alternatively, the homologous polypeptide may be one that is experimentally produced by directed mutation of one or more codons to alter the encoded amino acid of a CSP. In a preferred embodiment, the homologous polypeptide encodes a polypeptide that is a CSP.

Relatedness of proteins can also be characterized using a second functional test, such as the ability of a first protein competitively to inhibit the binding of a second protein to an antibody. It is, therefore, another aspect of the present invention to provide isolated polypeptides not only identical in sequence to those described with particularity herein, but also to provide isolated polypeptides ("cross-reactive proteins") that competitively inhibit the binding of antibodies to all or to a portion of the isolated polypeptides of the present invention. Such competitive inhibition can readily be determined using immunoassays well known in the art.

As discussed above, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes, and the sequence determined from one individual of a species may differ from other allelic forms present within the population. Thus, polypeptides of the present invention are also inclusive of those encoded by an allelic variant of a nucleic acid molecule encoding a CSP. In this embodiment, it is preferred that the polypeptide be encoded by an allelic variant of a gene that encodes a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO: 113-259. More preferred is that the polypeptide be encoded by an allelic variant of a gene that has the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1-112.

Polypeptides of the present invention are also inclusive of derivative polypeptides encoded by a nucleic acid molecule according to the instant invention. In this embodiment, it is preferred that the polypeptide be a CSP. Also preferred are derivative polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NO: 113-259 and which has been acetylated, carboxylated, phosphorylated, glycosylated, ubiquitinated or post-translationally modified in another manner. In another preferred embodiment, the derivative has been labeled with, *e.g.*, radioactive isotopes such as ^{125}I , ^{32}P , ^{35}S , and ^3H . In another preferred embodiment, the derivative has been labeled with fluorophores, chemiluminescent agents, enzymes, and antiligands that can serve as specific binding pair members for a labeled ligand.

Polypeptide modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance Creighton, Protein Structure and Molecular Properties, 2nd ed., W. H. Freeman and Company (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, in Johnson (ed.), Posttranslational Covalent Modification of Proteins, pgs. 1-12, Academic Press (1983); Seifter *et al.*, *Meth. Enzymol.* 182: 626-646 (1990) and Rattan *et al.*, *Ann. N.Y. Acad. Sci.* 663: 48-62 (1992).

One may determine whether a polypeptide of the invention is likely to be post-translationally modified by analyzing the sequence of the polypeptide to determine if there are peptide motifs indicative of sites for post-translational modification. There are a number of computer programs that permit prediction of post-translational modifications.

See, e.g., expasy.org (accessed November 11, 2002) of the world wide web, which includes PSORT, for prediction of protein sorting signals and localization sites, SignalP, for prediction of signal peptide cleavage sites, MITOPROT and Predotar, for prediction of mitochondrial targeting sequences, NetOGlyc, for prediction of type O-glycosylation sites
 5 in mammalian proteins, big-PI Predictor and DGPI, for prediction of prenylation-anchor and cleavage sites, and NetPhos, for prediction of Ser, Thr and Tyr phosphorylation sites in eukaryotic proteins. Other computer programs, such as those included in GCG, also may be used to determine post-translational modification peptide motifs.

General examples of types of post-translational modifications include, but are not
 10 limited to: (Z)-dehydrobutyrine; 1-chondroitin sulfate-L-aspartic acid ester; 1'-glycosyl-L-tryptophan; 1'-phospho-L-histidine; 1-thioglycine; 2'-(S-L-cysteinyl)-L-histidine; 2'-[3-carboxamido (trimethylammonio)propyl]-L-histidine; 2'-alpha-mannosyl-L-tryptophan; 2-methyl-L-glutamine; 2-oxobutanoic acid; 2-pyrrolidone carboxylic acid; 3'-(1'-L-histidyl)-L-tyrosine; 3'-(8alpha-FAD)-L-histidine; 3'-(S-L-cysteinyl)-L-tyrosine; 3', 3'',5'-triiodo-L-
 15 thyronine; 3'-4'-phospho-L-tyrosine; 3-hydroxy-L-proline; 3'-methyl-L-histidine; 3-methyl-L-lanthionine; 3'-phospho-L-histidine; 4'-(L-tryptophan)-L-tryptophyl quinone; 42 N-cysteinyl-glycosylphosphatidylinositoethanolamine; 43 -(T-L-histidyl)-L-tyrosine; 4-hydroxy-L-arginine; 4-hydroxy-L-lysine; 4-hydroxy-L-proline; 5'-(N6-L-lysine)-L-topaquinone; 5-hydroxy-L-lysine; 5-methyl-L-arginine; alpha-I-microglobulin-Ig alpha
 20 complex chromophore; bis-L-cysteinyl bis-L-histidino diiron disulfide; bis-L--cysteinyl-L-N3'-histidino-L-serinyl tetrairon' tetrasulfide; chondroitin sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-alanine; D-allo-isoleucine; D-asparagine; dehydroalanine; dehydrotyrosine; dermatan 4-sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-glucuronyl-N-glycine; dipyrrolylmethanemethyl-L-
 25 cysteine; D-leucine; D-methionine; D-phenylalanine; D-serine; D-tryptophan; glycine amide; glycine oxazolecarboxylic acid; glycine thiazolecarboxylic acid; heme P450-bis-L-cysteine-L-tyrosine; heme-bis-L-cysteine; hemediol-L-aspartyl ester-L-glutamyl ester; hemediol-L-aspartyl ester-L-glutamyl ester-L-methionine sulfonium; heme-L-cysteine; heme-L-histidine; heparan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-
 30 serine; heme P450-bis-L-cysteine-L-lysine; hexakis-L-cysteinyl hexairon hexasulfide; keratan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-threonine; L-oxoalanine- lactic acid; L phenyllactic acid; L'-(8alpha-FAD)-L-histidine; L-2'.4'.5'-topaquinone; L-3'.4'.5'-dihydroxyphenylalanine; L-3'.4'.5'-trihydroxyphenylalanine; L-4'-

- bromophenylalanine; L-6'-bromotryptophan; L-alanine amide; L-alanyl imidazolinone glycine; L-allysine; L-arginine amide; L-asparagine amide; L-aspartic 4-phosphoric anhydride; L-aspartic acid 1-amide; L-beta-methylthioaspartic acid; L-bromohistidine; L-citrulline; L-cysteine amide; L-cysteine glutathione disulfide; L-cysteine methyl disulfide;
- 5 L-cysteine methyl ester; L-cysteine oxazolecarboxylic acid; L-cysteine oxazolinecarboxylic acid; L-cysteine persulfide; L-cysteine sulfenic acid; L-cysteine sulfinic acid; L-cysteine thiazolecarboxylic acid; L-cysteinyl homocitryl molybdenum-heptairon-nonasulfide; L-cysteinyl imidazolinone glycine; L-cysteinyl molybdopterin; L-cysteinyl molybdopterin guanine dinucleotide; L-cystine; L-erythro-beta-
- 10 hydroxyasparagine; L-erythro-beta-hydroxyaspartic acid; L-gamma-carboxyglutamic acid; L-glutamic acid 1-amide; L-glutamic acid 5-methyl ester; L-glutamine amide; L-glutamyl 5-glycerylphosphorylethanolamine; L-histidine amide; L-isoglutamyl-polyglutamic acid; L-isoglutamyl-polyglycine; L-isoleucine amide; L-lanthionine; L-leucine amide; L-lysine amide; L-lysine thiazolecarboxylic acid; L-lysinoalanine; L-methionine amide; L-
- 15 methionine sulfone; L-phenylalanine thiazolecarboxylic acid; L-phenylalanine amide; L-proline amide; L-selenocysteine; L-selenocysteinyl molybdopterin guanine dinucleotide; L-serine amide; L-serine thiazolecarboxylic acid; L-seryl imidazolinone glycine; L-T-bromophenylalanine; L-T-bromophenylalanine; L-threonine amide; L-thyroxine; L-tryptophan amide; L-tryptophyl quinone; L-tyrosine amide; L-valine amide; meso-
- 20 lanthionine; N-(L-glutamyl)-L-tyrosine; N-(L-isoaspartyl)-glycine; N-(L-isoaspartyl)-L-cysteine; N,N,N-trimethyl-L-alanine; N,N-dimethyl-L-proline; N2-acetyl-L-lysine; N2-succinyl-L-tryptophan; N4-(ADP-ribosyl)-L-asparagine; N4-glycosyl-L-asparagine; N4-hydroxymethyl-L-asparagine; N4-methyl-L-asparagine; N5-methyl-L-glutamine; N6-1-carboxyethyl-L-lysine; N6-(4-amino hydroxybutyl)-L-lysine; N6-(L-isoglutamyl)-L-
- 25 lysine; N6-(phospho-5'-adenosine)-L-lysine; N6-(phospho-5'-guanosine)-L-lysine; N6,N6,N6-trimethyl-L-lysine; N6,N6-dimethyl-L-lysine; N6-acetyl-L-lysine; N6-biotinyl-L-lysine; N6-carboxy-L-lysine; N6-formyl-L-lysine; N6-glycyl-L-lysine; N6-lipoyl-L-lysine; N6-methyl-L-lysine; N6-methyl-N6-poly(N-methyl-propylamine)-L-lysine; N6-mureinyl-L-lysine; N6-myristoyl-L-lysine; N6-palmitoyl-L-lysine; N6-pyridoxal
- 30 phosphate-L-lysine; N6-pyruvic acid 2-iminyl-L-lysine; N6-retinal-L-lysine; N-acetyl-glycine; N-acetyl-L-glutamine; N-acetyl-L-alanine; N-acetyl-L-aspartic acid; N-acetyl-L-cysteine; N-acetyl-L-glutamic acid; N-acetyl-L-isoleucine; N-acetyl-L-methionine; N-acetyl-L-proline; N-acetyl-L-serine; N-acetyl-L-threonine; N-acetyl-L-

tyrosine; N-acetyl-L-valine; N-alanyl-glycosylphosphatidylinositoethanolamine; N-asparaginyl-glycosylphosphatidylinositoethanolamine; N-aspartyl-glycosylphosphatidylinositoethanolamine; N-formylglycine; N-formyl-L-methionine; N-glycyl-glycosylphosphatidylinositoethanolamine; N-L-glutamyl-poly-L-glutamic acid; N-methylglycine; N-methyl-L-alanine; N-methyl-L-methionine; N-methyl-L-phenylalanine;

5 N-myristoyl-glycine; N-palmitoyl-L-cysteine; N-pyruvic acid 2-iminyl-L-cysteine; N-pyruvic acid 2-iminyl-L-valine; N-seryl-glycosylphosphatidylinositoethanolamine; N-seryl-glycosylphosphatidylinositoethanolamine; O-(ADP-ribosyl)-L-serine; O-(phospho-5'-adenosine)-L-threonine; O-(phospho-5'-DNA)-L-serine; O-(phospho-5'-DNA)-L-threonine; O-(phospho-5'rRNA)-L-serine; O-(phosphoribosyl dephospho-coenzyme A)-L-serine; O-(sn-1-glycerophosphoryl)-L-serine; O4'-(8alpha-FAD)-L-tyrosine; O4'-(phospho-5'-adenosine)-L-tyrosine; O4'-(phospho-5'-DNA)-L-tyrosine; O4'-(phospho-5'-RNA)-L-tyrosine; O4'-(phospho-5'-uridine)-L-tyrosine; O4-glycosyl-L-hydroxyproline; O4-glycosyl-L-tyrosine; O4'-sulfo-L-tyrosine; O5-glycosyl-L-hydroxylysine; O-glycosyl-L-serine; O-glycosyl-L-threonine; omega-N-(ADP-ribosyl)-L-arginine; omega-N-omega-N'-dimethyl-L-arginine; omega-N-methyl-L-arginine; omega-N-omega-N-dimethyl-L-arginine; omega-N-phospho-L-arginine; O'octanoyl-L-serine; O-palmitoyl-L-serine; O-palmitoyl-L-threonine; O-phospho-L-serine; O-phospho-L-threonine; O-phosphopantetheine-L-serine; phycoerythrobilin-bis-L-cysteine; phycourobilin-bis-L-cysteine;

20 pyrroloquinoline quinone; pyruvic acid; S hydroxycinnamyl-L-cysteine; S-(2-aminovinyl) methyl-D-cysteine; S-(2-aminovinyl)-D-cysteine; S-(6-FW-L-cysteine; S-(8alpha-FAD)-L-cysteine; S-(ADP-ribosyl)-L-cysteine; S-(L-isoglutamyl)-L-cysteine; S-12-hydroxyfarnesyl-L-cysteine; S-acetyl-L-cysteine; S-diacylglycerol-L-cysteine; S-diphytanylglycerol diether-L-cysteine; S-farnesyl-L-cysteine; S-geranylgeranyl-L-cysteine; S-glycosyl-L-cysteine; S-glycyl-L-cysteine; S-methyl-L-cysteine; S-nitrosyl-L-cysteine; S-palmitoyl-L-cysteine; S-phospho-L-cysteine; S-phycobiliviolin-L-cysteine; S-phycocyanobilin-L-cysteine; S-phycoerythrobilin-L-cysteine; S-phytochromobilin-L-cysteine; S-selenyl-L-cysteine; S-sulfo-L-cysteine; tetrakis-L-cysteiny l diiron disulfide; tetrakis-L-cysteiny l iron; tetrakis-L-cysteiny l tetrairon tetrasulfide; trans-2,3-cis 4-dihydroxy-L-proline; tris-L-cysteiny l triiron tetrasulfide; tris-L-cysteiny l triiron trisulfide; tris-L-cysteiny l-L-aspartato tetrairon tetrasulfide; tris-L-cysteiny l-L-cysteine persulfido-bis-L-glutamato-L-histidino tetrairon disulfide trioxide; tris-L-cysteiny l-L-N3'-histidino

30

tetrairon tetrasulfide; tris-L-cysteinyl-L-NI'-histidino tetrairon tetrasulfide; and tris-L-cysteinyl-L-serinyl tetrairon tetrasulfide.

Additional examples of PTMs may be found in web sites such as the Delta Mass database based on Krishna, R. G. and F. Wold (1998). Posttranslational Modifications.

5 Proteins - Analysis and Design. R. H. Angeletti. San Diego, Academic Press. 1: 121-206; Methods in Enzymology, 193, J.A. McClosky (ed) (1990), pages 647-660; Methods in Protein Sequence Analysis edited by Kazutomo Imahori and Fumio Sakiyama, Plenum Press, (1993) "Post-translational modifications of proteins" R.G. Krishna and F. Wold pages 167-172; "GlycoSuiteDB: a new curated relational database of glycoprotein glycan

10 structures and their biological sources" Cooper et al. Nucleic Acids Res. 29; 332-335 (2001) "O-GLYCBASE version 4.0: a revised database of O-glycosylated proteins" Gupta et al. Nucleic Acids Research, 27: 370-372 (1999); and "PhosphoBase, a database of phosphorylation sites: release 2.0.", Kreegipuu et al. Nucleic Acids Res 27(1):237-239 (1999) see also, WO 02/21139A2, the disclosure of which is incorporated herein by

15 reference in its entirety.

Tumorigenesis is often accompanied by alterations in the post-translational modifications of proteins. Thus, in another embodiment, the invention provides polypeptides from cancerous cells or tissues that have altered post-translational modifications compared to the post-translational modifications of polypeptides from

20 normal cells or tissues. A number of altered post-translational modifications are known. One common alteration is a change in phosphorylation state, wherein the polypeptide from the cancerous cell or tissue is hyperphosphorylated or hypophosphorylated compared to the polypeptide from a normal tissue, or wherein the polypeptide is phosphorylated on different residues than the polypeptide from a normal cell. Another common alteration is

25 a change in glycosylation state, wherein the polypeptide from the cancerous cell or tissue has more or less glycosylation than the polypeptide from a normal tissue, and/or wherein the polypeptide from the cancerous cell or tissue has a different type of glycosylation than the polypeptide from a noncancerous cell or tissue. Changes in glycosylation may be critical because carbohydrate-protein and carbohydrate-carbohydrate interactions are

30 important in cancer cell progression, dissemination and invasion. See, e.g., Barchi, *Curr. Pharm. Des.* 6: 485-501 (2000), Verma, *Cancer Biochem. Biophys.* 14: 151-162 (1994) and Dennis et al., *Bioessays* 5: 412-421 (1999).

Another post-translational modification that may be altered in cancer cells is prenylation. Prenylation is the covalent attachment of a hydrophobic prenyl group (either farnesyl or geranylgeranyl) to a polypeptide. Prenylation is required for localizing a protein to a cell membrane and is often required for polypeptide function. For instance,
5 the Ras superfamily of GTPase signalling proteins must be prenylated for function in a cell. See, e.g., Prendergast et al., *Semin. Cancer Biol.* 10: 443-452 (2000) and Khwaja et al., *Lancet* 355: 741-744 (2000).

Other post-translation modifications that may be altered in cancer cells include, without limitation, polypeptide methylation, acetylation, arginylation or racemization of
10 amino acid residues. In these cases, the polypeptide from the cancerous cell may exhibit either increased or decreased amounts of the post-translational modification compared to the corresponding polypeptides from noncancerous cells.

Other polypeptide alterations in cancer cells include abnormal polypeptide cleavage of proteins and aberrant protein-protein interactions. Abnormal polypeptide
15 cleavage may be cleavage of a polypeptide in a cancerous cell that does not usually occur in a normal cell, or a lack of cleavage in a cancerous cell, wherein the polypeptide is cleaved in a normal cell. Aberrant protein-protein interactions may be either covalent cross-linking or non-covalent binding between proteins that do not normally bind to each other. Alternatively, in a cancerous cell, a protein may fail to bind to another protein to
20 which it is bound in a noncancerous cell. Alterations in cleavage or in protein-protein interactions may be due to over- or underproduction of a polypeptide in a cancerous cell compared to that in a normal cell, or may be due to alterations in post-translational modifications (see above) of one or more proteins in the cancerous cell. See, e.g., Henschen-Edman, *Ann. N.Y. Acad. Sci.* 936: 580-593 (2001).

25 Alterations in polypeptide post-translational modifications, as well as changes in polypeptide cleavage and protein-protein interactions, may be determined by any method known in the art. For instance, alterations in phosphorylation may be determined by using anti-phosphoserine, anti-phosphothreonine or anti-phosphotyrosine antibodies or by amino acid analysis. Glycosylation alterations may be determined using antibodies specific for
30 different sugar residues, by carbohydrate sequencing, or by alterations in the size of the glycoprotein, which can be determined by, e.g., SDS polyacrylamide gel electrophoresis (PAGE). Other alterations of post-translational modifications, such as prenylation, racemization, methylation, acetylation and arginylation, may be determined by chemical

analysis, protein sequencing, amino acid analysis, or by using antibodies specific for the particular post-translational modifications. Changes in protein-protein interactions and in polypeptide cleavage may be analyzed by any method known in the art including, without limitation, non-denaturing PAGE (for non-covalent protein-protein interactions), SDS
5 PAGE (for covalent protein-protein interactions and protein cleavage), chemical cleavage, protein sequencing or immunoassays.

In another embodiment, the invention provides polypeptides that have been post-translationally modified. In one embodiment, polypeptides may be modified enzymatically or chemically, by addition or removal of a post-translational modification.
10 For example, a polypeptide may be glycosylated or deglycosylated enzymatically. Similarly, polypeptides may be phosphorylated using a purified kinase, such as a MAP kinase (e.g., p38, ERK, or JNK) or a tyrosine kinase (e.g., Src or erbB2). A polypeptide may also be modified through synthetic chemistry. Alternatively, one may isolate the polypeptide of interest from a cell or tissue that expresses the polypeptide with the desired
15 post-translational modification. In another embodiment, a nucleic acid molecule encoding the polypeptide of interest is introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide in the desired fashion. If the polypeptide does not contain a motif for a desired post-translational modification, one may alter the post-translational modification by mutating the nucleic acid sequence of a nucleic
20 acid molecule encoding the polypeptide so that it contains a site for the desired post-translational modification. Amino acid sequences that may be post-translationally modified are known in the art. See, e.g., the programs described above on the website expasy.org of the world wide web. The nucleic acid molecule may also be introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide.
25 Similarly, one may delete sites that are post-translationally modified by either mutating the nucleic acid sequence so that the encoded polypeptide does not contain the post-translational modification motif, or by introducing the native nucleic acid molecule into a host cell that is not capable of post-translationally modifying the encoded polypeptide.

It will be appreciated, as is well known and as noted above, that polypeptides are
30 not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing events and events brought about by human manipulation which do not occur naturally. Circular, branched and branched

circular polypeptides may be synthesized by non-translation natural processes and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

Useful post-synthetic (and post-translational) modifications include conjugation to detectable labels, such as fluorophores. A wide variety of amine-reactive and thiol-reactive fluorophore derivatives have been synthesized that react under nondenaturing conditions with N-terminal amino groups and epsilon amino groups of lysine residues, on the one hand, and with free thiol groups of cysteine residues, on the other.

Kits are available commercially that permit conjugation of proteins to a variety of amine-reactive or thiol-reactive fluorophores: Molecular Probes, Inc. (Eugene, OR, USA), *e.g.*, offers kits for conjugating proteins to Alexa Fluor 350, Alexa Fluor 430, Fluorescein-EX, Alexa Fluor 488, Oregon Green 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 568, Alexa Fluor 594, and Texas Red-X.

A wide variety of other amine-reactive and thiol-reactive fluorophores are available commercially (Molecular Probes, Inc., Eugene, OR, USA), including Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, OR, USA).

The polypeptides of the present invention can also be conjugated to fluorophores, other proteins, and other macromolecules, using bifunctional linking reagents. Common homobifunctional reagents include, *e.g.*, APG, AEDP, BASED, BMB, BMDB, BMH, BMOE, BM[PEO]3, BM[PEO]4, BS3, BSOCOES, DFDNB, DMA, DMP, DMS, DPDPB,

DSG, DSP (Lomant's Reagent), DSS, DST, DTBP, DTME, DTSSP, EGS, HBVS, Sulfo-BSOCOES, Sulfo-DST, Sulfo-EGS (all available from Pierce, Rockford, IL, USA); common heterobifunctional cross-linkers include ABH, AMAS, ANB-NOS, APDP, ASBA, BMPA, BMPH, BMPS, EDC, EMCA, EMCH, EMCS, KMUA, KMUH, GMBS, LC-SMCC, LC-SPDP, MBS, M2C2H, MPBH, MSA, NHS-ASA, PDPH, PMPI, SADP, SAED, SAND, SANPAH, SASD, SATP, SBAP, SFAD, SIA, SIAB, SMCC, SMPB, SMPH, SMPT, SPDP, Sulfo-EMCS, Sulfo-GMBS, Sulfo-HSAB, Sulfo-KMUS, Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH, Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB, Sulfo-LC-SMPT, SVSB, TFCS (all available
5
10 Pierce, Rockford, IL, USA).

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can be conjugated, using such cross-linking reagents, to fluorophores that are not amine- or thiol-reactive. Other labels that usefully can be conjugated to polypeptides of the present invention include radioactive labels, echosonographic contrast reagents, and MRI contrast agents.
15

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can also usefully be conjugated using cross-linking agents to carrier proteins, such as KLH, bovine thyroglobulin, and even bovine serum albumin (BSA), to increase immunogenicity for raising anti-CSP antibodies.

20 Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can also usefully be conjugated to polyethylene glycol (PEG); PEGylation increases the serum half life of proteins administered intravenously for replacement therapy. Delgado *et al.*, *Crit. Rev. Ther. Drug Carrier Syst.* 9(3-4): 249-304 (1992); Scott *et al.*, *Curr. Pharm. Des.* 4(6): 423-38 (1998); DeSantis *et al.*, *Curr. Opin. Biotechnol.* 10(4): 324-30 (1999). PEG monomers can be attached to the protein directly or through a linker, with PEGylation using PEG monomers activated with tresyl chloride (2,2,2-trifluoroethanesulphonyl chloride) permitting direct attachment under mild conditions.
25

Polypeptides of the present invention are also inclusive of analogs of a polypeptide encoded by a nucleic acid molecule according to the instant invention. In a preferred
30 embodiment, this polypeptide is a CSP. In a more preferred embodiment, this polypeptide is derived from a polypeptide having part or all of the amino acid sequence of SEQ ID NO: 113-259. Also preferred is an analog polypeptide comprising one or more

substitutions of non-natural amino acids or non-native inter-residue bonds compared to the naturally occurring polypeptide. In one embodiment, the analog is structurally similar to a CSP, but one or more peptide linkages is replaced by a linkage selected from the group consisting of --CH₂NH--, --CH₂S--, --CH₂-CH₂--, --CH=CH--(cis and trans), --COCH₂--,
5 --CH(OH)CH₂-- and --CH₂SO--. In another embodiment, the analog comprises substitution of one or more amino acids of a CSP with a D-amino acid of the same type or other non-natural amino acid in order to generate more stable peptides. D-amino acids can readily be incorporated during chemical peptide synthesis: peptides assembled from D-amino acids are more resistant to proteolytic attack; incorporation of D-amino acids can
10 also be used to confer specific three-dimensional conformations on the peptide. Other amino acid analogues commonly added during chemical synthesis include ornithine, norleucine, phosphorylated amino acids (typically phosphoserine, phosphothreonine, phosphotyrosine), L-malonyltyrosine, a non-hydrolyzable analog of phosphotyrosine (*see, e.g., Krole et al., Biochem. Biophys. Res. Com.* 209: 817-821 (1995)), and various
15 halogenated phenylalanine derivatives.

Non-natural amino acids can be incorporated during solid phase chemical synthesis or by recombinant techniques, although the former is typically more common. Solid phase chemical synthesis of peptides is well established in the art. Procedures are described, *inter alia*, in Chan *et al.* (eds.), Fmoc Solid Phase Peptide Synthesis: A
20 Practical Approach (Practical Approach Series), Oxford Univ. Press (March 2000); Jones, Amino Acid and Peptide Synthesis (Oxford Chemistry Primers, No 7), Oxford Univ. Press (1992); and Bodanszky, Principles of Peptide Synthesis (Springer Laboratory), Springer Verlag (1993).

Amino acid analogues having detectable labels are also usefully incorporated
25 during synthesis to provide derivatives and analogs. Biotin, for example can be added using biotinoyl-(9-fluorenylmethoxycarbonyl)-L-lysine (Fmoc biocytin) (Molecular Probes, Eugene, OR, USA). Biotin can also be added enzymatically by incorporation into a fusion protein of an *E. coli* BirA substrate peptide. The Fmoc and tBOC derivatives of dabcyL-L-lysine (Molecular Probes, Inc., Eugene, OR, USA) can be used to incorporate
30 the dabcyL chromophore at selected sites in the peptide sequence during synthesis. The aminonaphthalene derivative EDANS, the most common fluorophore for pairing with the dabcyL quencher in fluorescence resonance energy transfer (FRET) systems, can be

introduced during automated synthesis of peptides by using EDANS-FMOC-L-glutamic acid or the corresponding *t*BOC derivative (both from Molecular Probes, Inc., Eugene, OR, USA). Tetramethylrhodamine fluorophores can be incorporated during automated FMOC synthesis of peptides using (FMOC)-TMR-L-lysine (Molecular Probes, Inc.

5 Eugene, OR, USA).

Other useful amino acid analogues that can be incorporated during chemical synthesis include aspartic acid, glutamic acid, lysine, and tyrosine analogues having allyl side-chain protection (Applied Biosystems, Inc., Foster City, CA, USA); the allyl side chain permits synthesis of cyclic, branched-chain, sulfonated, glycosylated, and

10 phosphorylated peptides.

A large number of other FMOC-protected non-natural amino acid analogues capable of incorporation during chemical synthesis are available commercially, including, *e.g.*, Fmoc-2-aminobicyclo[2.2.1]heptane-2-carboxylic acid, Fmoc-3-endo-aminobicyclo[2.2.1]heptane-2-endo-carboxylic acid, Fmoc-3-exo-aminobicyclo[2.2.1]heptane-2-exo-carboxylic acid, Fmoc-3-endo-amino-bicyclo[2.2.1]hept-5-ene-2-endo-carboxylic acid, Fmoc-3-exo-amino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid, Fmoc-cis-2-amino-1-cyclohexanecarboxylic acid, Fmoc-trans-2-amino-1-cyclohexanecarboxylic acid, Fmoc-1-amino-1-cyclopentanecarboxylic acid, Fmoc-cis-2-amino-1-cyclopentanecarboxylic acid, Fmoc-1-amino-1-cyclopropanecarboxylic acid, Fmoc-D-2-amino-4-(ethylthio)butyric acid, Fmoc-L-2-amino-4-(ethylthio)butyric acid, Fmoc-L-buthionine, Fmoc-S-methyl-L-Cysteine, Fmoc-2-aminobenzoic acid (anthranillic acid), Fmoc-3-aminobenzoic acid, Fmoc-4-aminobenzoic acid, Fmoc-2-aminobenzophenone-2'-carboxylic acid, Fmoc-N-(4-aminobenzoyl)- β -alanine, Fmoc-2-amino-4,5-dimethoxybenzoic acid, Fmoc-4-aminohippuric acid, Fmoc-2-amino-3-hydroxybenzoic acid, Fmoc-2-amino-5-hydroxybenzoic acid, Fmoc-3-amino-4-hydroxybenzoic acid, Fmoc-4-amino-3-hydroxybenzoic acid, Fmoc-4-amino-2-hydroxybenzoic acid, Fmoc-5-amino-2-hydroxybenzoic acid, Fmoc-2-amino-3-methoxybenzoic acid, Fmoc-4-amino-3-methoxybenzoic acid, Fmoc-2-amino-3-methylbenzoic acid, Fmoc-2-amino-5-methylbenzoic acid, Fmoc-2-amino-6-methylbenzoic acid, Fmoc-3-amino-2-methylbenzoic acid, Fmoc-3-amino-4-methylbenzoic acid, Fmoc-4-amino-3-methylbenzoic acid, Fmoc-3-amino-2-naphtoic acid, Fmoc-D,L-3-amino-3-phenylpropionic acid, Fmoc-L-Methyldopa, Fmoc-2-amino-4,6-dimethyl-3-

pyridinecarboxylic acid, Fmoc-D,L-amino-2-thiophenacetic acid, Fmoc-4-(carboxymethyl)piperazine, Fmoc-4-carboxypiperazine, Fmoc-4-(carboxymethyl)homopiperazine, Fmoc-4-phenyl-4-piperidinecarboxylic acid, Fmoc-L-1,2,3,4-tetrahydronorharman-3-carboxylic acid, Fmoc-L-thiazolidine-4-carboxylic acid, all
5 available from The Peptide Laboratory (Richmond, CA, USA).

Non-natural residues can also be added biosynthetically by engineering a suppressor tRNA, typically one that recognizes the UAG stop codon, by chemical aminoacylation with the desired unnatural amino acid. Conventional site-directed mutagenesis is used to introduce the chosen stop codon UAG at the site of interest in the
10 protein gene. When the acylated suppressor tRNA and the mutant gene are combined in an *in vitro* transcription/translation system, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing that amino acid at the specified position. Liu *et al.*, *Proc. Natl Acad. Sci. USA* 96(9): 4780-5 (1999); Wang *et al.*, *Science* 292(5516): 498-500 (2001).

15 *Fusion Proteins*

Another aspect of the present invention relates to the fusion of a polypeptide of the present invention to heterologous polypeptides. In a preferred embodiment, the polypeptide of the present invention is a CSP. In a more preferred embodiment, the polypeptide of the present invention that is fused to a heterologous polypeptide which
20 comprises part or all of the amino acid sequence of SEQ ID NO: 113-259, or is a mutein, homologous polypeptide, analog or derivative thereof. In an even more preferred embodiment, the fusion protein is encoded by a nucleic acid molecule comprising all or part of the nucleic acid sequence of SEQ ID NO: 1-112, or comprises all or part of a nucleic acid sequence that selectively hybridizes or is homologous to a nucleic acid
25 molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112.

The fusion proteins of the present invention will include at least one fragment of a polypeptide of the present invention, which fragment is at least 6, typically at least 8, often at least 15, and usefully at least 16, 17, 18, 19, or 20 amino acids long. The fragment of the polypeptide of the present to be included in the fusion can usefully be at least 25
30 amino acids long, at least 50 amino acids long, and can be at least 75, 100, or even 150 amino acids long. Fusions that include the entirety of a polypeptide of the present invention have particular utility.

The heterologous polypeptide included within the fusion protein of the present invention is at least 6 amino acids in length, often at least 8 amino acids in length, and preferably at least 15, 20, or 25 amino acids in length. Fusions that include larger polypeptides, such as the IgG Fc region, and even entire proteins (such as GFP
5 chromophore-containing proteins) are particularly useful.

As described above in the description of vectors and expression vectors of the present invention, which discussion is incorporated here by reference in its entirety, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those designed to facilitate purification and/or visualization of
10 recombinantly-expressed proteins. *See, e.g.*, Ausubel, Chapter 16, (1992), *supra*. Although purification tags can also be incorporated into fusions that are chemically synthesized, chemical synthesis typically provides sufficient purity that further purification by HPLC suffices; however, visualization tags as above described retain their utility even when the protein is produced by chemical synthesis, and when so included
15 render the fusion proteins of the present invention useful as directly detectable markers of the presence of a polypeptide of the invention.

As also discussed above, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those that facilitate secretion of recombinantly expressed proteins into the periplasmic space or extracellular milieu for
20 prokaryotic hosts or into the culture medium for eukaryotic cells through incorporation of secretion signals and/or leader sequences. For example, a His⁶ tagged protein can be purified on a Ni affinity column and a GST fusion protein can be purified on a glutathione affinity column. Similarly, a fusion protein comprising the Fc domain of IgG can be purified on a Protein A or Protein G column and a fusion protein comprising an epitope
25 tag such as myc can be purified using an immunoaffinity column containing an anti-c-myc antibody. It is preferable that the epitope tag be separated from the protein encoded by the essential gene by an enzymatic cleavage site that can be cleaved after purification. See also the discussion of nucleic acid molecules encoding fusion proteins that may be expressed on the surface of a cell.

30 Other useful fusion proteins of the present invention include those that permit use of the polypeptide of the present invention as bait in a yeast two-hybrid system. *See* Bartel *et al.* (eds.), The Yeast Two-Hybrid System, Oxford University Press (1997); Zhu *et al.*, Yeast Hybrid Technologies, Eaton Publishing (2000); Fields *et al.*, *Trends Genet.*

10(8): 286-92 (1994); Mendelsohn *et al.*, *Curr. Opin. Biotechnol.* 5(5): 482-6 (1994); Luban *et al.*, *Curr. Opin. Biotechnol.* 6(1): 59-64 (1995); Allen *et al.*, *Trends Biochem. Sci.* 20(12): 511-6 (1995); Drees, *Curr. Opin. Chem. Biol.* 3(1): 64-70 (1999); Topcu *et al.*, *Pharm. Res.* 17(9): 1049-55 (2000); Fashena *et al.*, *Gene* 250(1-2): 1-14 (2000); Colas
5 *et al.*, *Nature* 380, 548-550 (1996); Norman, T. *et al.*, *Science* 285, 591-595 (1999); Fabbri *et al.*, *Oncogene* 18, 4357-4363 (1999); Xu *et al.*, *Proc Natl Acad Sci U S A.* 94, 12473-12478 (1997); Yang, *et al.*, *Nuc. Acids Res.* 23, 1152-1156 (1995); Kolonin *et al.*, *Proc Natl Acad Sci U S A* 95, 14266-14271 (1998); Cohen *et al.*, *Proc Natl Acad Sci U S A* 95, 14272-14277 (1998); Uetz, *et al.* *Nature* 403, 623-627(2000); Ito, *et al.*, *Proc Natl Acad Sci U S A* 98, 4569-4574 (2001). Typically, such fusion is to either *E. coli* LexA or yeast GAL4 DNA binding domains. Related bait plasmids are available that express the bait fused to a nuclear localization signal.

Other useful fusion proteins include those that permit display of the encoded polypeptide on the surface of a phage or cell, fusions to intrinsically fluorescent proteins,
15 such as green fluorescent protein (GFP), and fusions to the IgG Fc region, as described above.

The polypeptides of the present invention can also usefully be fused to protein toxins, such as Pseudomonas exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin, in order to effect ablation of cells that bind or take up the proteins of
20 the present invention.

Fusion partners include, *inter alia*, *myc*, hemagglutinin (HA), GST, immunoglobulins, β -galactosidase, biotin trpE, protein A, β -lactamase, α -amylase, maltose binding protein, alcohol dehydrogenase, polyhistidine (for example, six histidine at the amino and/or carboxyl terminus of the polypeptide), lacZ, green fluorescent protein
25 (GFP), yeast α mating factor, GAL4 transcription activation or DNA binding domain, luciferase, and serum proteins such as ovalbumin, albumin and the constant domain of IgG. *See, e.g.*, Ausubel (1992), *supra* and Ausubel (1999), *supra*. Fusion proteins may also contain sites for specific enzymatic cleavage, such as a site that is recognized by enzymes such as Factor XIII, trypsin, pepsin, or any other enzyme known in the art.
30 Fusion proteins will typically be made by either recombinant nucleic acid methods, as described above, chemically synthesized using techniques well known in the art (*e.g.*, a Merrifield synthesis), or produced by chemical cross-linking.

Another advantage of fusion proteins is that the epitope tag can be used to bind the fusion protein to a plate or column through an affinity linkage for screening binding proteins or other molecules that bind to the CSP.

As further described below, the polypeptides of the present invention can readily
5 be used as specific immunogens to raise antibodies that specifically recognize polypeptides of the present invention including CSPs and their allelic variants and homologues. The antibodies, in turn, can be used, *inter alia*, specifically to assay for the polypeptides of the present invention, particularly CSPs, *e.g.* by ELISA for detection of protein fluid samples, such as serum, by immunohistochemistry or laser scanning
10 cytometry, for detection of protein in tissue samples, or by flow cytometry, for detection of intracellular protein in cell suspensions, for specific antibody-mediated isolation and/or purification of CSPs, as for example by immunoprecipitation, and for use as specific agonists or antagonists of CSPs.

One may determine whether polypeptides of the present invention including CSPs,
15 muteins, homologous proteins or allelic variants or fusion proteins of the present invention are functional by methods known in the art. For instance, residues that are tolerant of change while retaining function can be identified by altering the polypeptide at known residues using methods known in the art, such as alanine scanning mutagenesis, Cunningham *et al.*, *Science* 244(4908): 1081-5 (1989); transposon linker scanning
20 mutagenesis, Chen *et al.*, *Gene* 263(1-2): 39-48 (2001); combinations of homolog- and alanine-scanning mutagenesis, Jin *et al.*, *J. Mol. Biol.* 226(3): 851-65 (1992); and combinatorial alanine scanning, Weiss *et al.*, *Proc. Natl. Acad. Sci USA* 97(16): 8950-4 (2000), followed by functional assay. Transposon linker scanning kits are available commercially (New England Biolabs, Beverly, MA, USA, catalog. no. E7-102S;
25 EZ::TN™ In-Frame Linker Insertion Kit, catalogue no. EZI04KN, (Epicentre Technologies Corporation, Madison, WI, USA).

Purification of the polypeptides or fusion proteins of the present invention is well known and within the skill of one having ordinary skill in the art. *See, e.g.*, Scopes, Protein Purification, 2d ed. (1987). Purification of recombinantly expressed polypeptides
30 is described above. Purification of chemically-synthesized peptides can readily be effected, *e.g.*, by HPLC.

Accordingly, it is an aspect of the present invention to provide the isolated polypeptides or fusion proteins of the present invention in pure or substantially pure form

in the presence or absence of a stabilizing agent. Stabilizing agents include both proteinaceous and non-proteinaceous material and are well known in the art. Stabilizing agents, such as albumin and polyethylene glycol (PEG) are known and are commercially available.

5 Although high levels of purity are preferred when the isolated polypeptide or fusion protein of the present invention are used as therapeutic agents, such as in vaccines and replacement therapy, the isolated polypeptides of the present invention are also useful at lower purity. For example, partially purified polypeptides of the present invention can be used as immunogens to raise antibodies in laboratory animals.

10 In a preferred embodiment, the purified and substantially purified polypeptides of the present invention are in compositions that lack detectable ampholytes, acrylamide monomers, bis-acrylamide monomers, and polyacrylamide.

 The polypeptides or fusion proteins of the present invention can usefully be attached to a substrate. The substrate can be porous or solid, planar or non-planar; the
15 bond can be covalent or noncovalent. For example, the peptides of the invention may be stabilized by covalent linkage to albumin. See, U.S. Patent No. 5,876,969, the contents of which are hereby incorporated in its entirety.

 The polypeptides or fusion proteins of the present invention can also be usefully bound to a porous substrate, commonly a membrane, typically comprising nitrocellulose, polyvinylidene fluoride (PVDF), or cationically derivatized, hydrophilic PVDF; so bound,
20 the polypeptides or fusion proteins of the present invention can be used to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized polypeptide or fusion protein of the present invention.

 As another example, the polypeptides or fusion proteins of the present invention
25 can usefully be bound to a substantially nonporous substrate, such as plastic, to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized protein of the present invention. Such plastics include polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate,
30 nitrocellulose, or mixtures thereof; when the assay is performed in a standard microtiter dish, the plastic is typically polystyrene.

 The polypeptides and fusion proteins of the present invention can also be attached to a substrate suitable for use as a surface enhanced laser desorption ionization source; so

attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biologic interaction there between.

5 The polypeptides or fusion proteins of the present invention can also be attached to a substrate suitable for use in surface plasmon resonance detection; so attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biological interaction there between.

Alternative Transcripts

10 In another aspect, the present invention provides splice variants of genes and proteins encoded thereby. The identification of a novel splice variant which encodes an amino acid sequence with a novel region can be targeted for the generation of reagents for use in detection and/or treatment of cancer. The novel amino acid sequence may lead to a unique protein structure, protein subcellular localization, biochemical processing or
15 function of the splice variant. This information can be used to directly or indirectly facilitate the generation of additional or novel therapeutics or diagnostics. The nucleotide sequence in this novel splice variant can be used as a nucleic acid probe for the diagnosis and/or treatment of cancer.

Specifically, the newly identified sequences may enable the production of new
20 antibodies or compounds directed against the novel region for use as a therapeutic or diagnostic. Alternatively, the newly identified sequences may alter the biochemical or biological properties of the encoded protein in such a way as to enable the generation of improved or different therapeutics targeting this protein.

Antibodies

25 In another aspect, the invention provides antibodies, including fragments and derivatives thereof, that bind specifically to polypeptides encoded by the nucleic acid molecules of the invention. In a preferred embodiment, the antibodies are specific for a polypeptide that is a CSP, or a fragment, mutein, derivative, analog or fusion protein thereof. In a more preferred embodiment, the antibodies are specific for a polypeptide that
30 comprises SEQ ID NO: 113-259, or a fragment, mutein, derivative, analog or fusion protein thereof.

The antibodies of the present invention can be specific for linear epitopes, discontinuous epitopes, or conformational epitopes of such proteins or protein fragments, either as present on the protein in its native conformation or, in some cases, as present on the proteins as denatured, as, *e.g.*, by solubilization in SDS. New epitopes may also be
5 due to a difference in post translational modifications (PTMs) in disease versus normal tissue. For example, a particular site on a CSP may be glycosylated in cancerous cells, but not glycosylated in normal cells or vice versa. In addition, alternative splice forms of a CSP may be indicative of cancer. Differential degradation of the C or N-terminus of a CSP may also be a marker or target for anticancer therapy. For example, a CSP may be
10 N-terminal degraded in cancer cells exposing new epitopes to antibodies which may selectively bind for diagnostic or therapeutic uses.

As is well known in the art, the degree to which an antibody can discriminate among molecular species in a mixture will depend, in part, upon the conformational relatedness of the species in the mixture; typically, the antibodies of the present invention
15 will discriminate over adventitious binding to non-CSP polypeptides by at least two-fold, more typically by at least 5-fold, typically by more than 10-fold, 25-fold, 50-fold, 75-fold, and often by more than 100-fold, and on occasion by more than 500-fold or 1000-fold. When used to detect the proteins or protein fragments of the present invention, the antibody of the present invention is sufficiently specific when it can be used to determine
20 the presence of the polypeptide of the present invention in samples derived from human colon.

Typically, the affinity or avidity of an antibody (or antibody multimer, as in the case of an IgM pentamer) of the present invention for a protein or protein fragment of the present invention will be at least about 1×10^{-6} molar (M), typically at least about 5×10^{-7}
25 M, 1×10^{-7} M, with affinities and avidities of at least 1×10^{-8} M, 5×10^{-9} M, 1×10^{-10} M and up to 1×10^{-13} M proving especially useful.

The antibodies of the present invention can be naturally occurring forms, such as IgG, IgM, IgD, IgE, IgY, and IgA, from any avian, reptilian, or mammalian species.

Human antibodies can, but will infrequently, be drawn directly from human donors
30 or human cells. In such case, antibodies to the polypeptides of the present invention will typically have resulted from fortuitous immunization, such as autoimmune immunization, with the polypeptide of the present invention. Such antibodies will typically, but will not

invariably, be polyclonal. In addition, individual polyclonal antibodies may be isolated and cloned to generate monoclonals.

Human antibodies are more frequently obtained using transgenic animals that express human immunoglobulin genes, which transgenic animals can be affirmatively immunized with the protein immunogen of the present invention. Human Ig-transgenic mice capable of producing human antibodies and methods of producing human antibodies therefrom upon specific immunization are described, *inter alia*, in U.S. Patent Nos. 6,162,963; 6,150,584; 6,114,598; 6,075,181; 5,939,598; 5,877,397; 5,874,299; 5,814,318; 5,789,650; 5,770,429; 5,661,016; 5,633,425; 5,625,126; 5,569,825; 5,545,807; 5,545,806, and 5,591,669, the disclosures of which are incorporated herein by reference in their entireties. Such antibodies are typically monoclonal, and are typically produced using techniques developed for production of murine antibodies.

Human antibodies are particularly useful, and often preferred, when the antibodies of the present invention are to be administered to human beings as *in vivo* diagnostic or therapeutic agents, since recipient immune response to the administered antibody will often be substantially less than that occasioned by administration of an antibody derived from another species, such as mouse.

IgG, IgM, IgD, IgE, IgY, and IgA antibodies of the present invention are also usefully obtained from other species, including mammals such as rodents (typically mouse, but also rat, guinea pig, and hamster), lagomorphs (typically rabbits), and also larger mammals, such as sheep, goats, cows, and horses; or egg laying birds or reptiles such as chickens or alligators. In such cases, as with the transgenic human-antibody-producing non-human mammals, fortuitous immunization is not required, and the non-human mammal is typically affirmatively immunized, according to standard immunization protocols, with the polypeptide of the present invention. One form of avian antibodies may be generated using techniques described in WO 00/29444, published 25 May 2000, which is herein incorporated by reference in its entirety.

As discussed above, virtually all fragments of 8 or more contiguous amino acids of a polypeptide of the present invention can be used effectively as immunogens when conjugated to a carrier, typically a protein such as bovine thyroglobulin, keyhole limpet hemocyanin, or bovine serum albumin, conveniently using a bifunctional linker such as those described elsewhere above, which discussion is incorporated by reference here.

Immunogenicity can also be conferred by fusion of the polypeptides of the present invention to other moieties. For example, polypeptides of the present invention can be produced by solid phase synthesis on a branched polylysine core matrix; these multiple antigenic peptides (MAPs) provide high purity, increased avidity, accurate chemical definition and improved safety in vaccine development. Tam *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 5409-5413 (1988); Posnett *et al.*, *J. Biol. Chem.* 263: 1719-1725 (1988).

Protocols for immunizing non-human mammals or avian species are well-established in the art. See Harlow *et al.* (eds.), Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998); Coligan *et al.* (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc. (2001); Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000); Gross M, Speck *J.Dtsch. Tierarztl. Wochenschr.* 103: 417-422 (1996). Immunization protocols often include multiple immunizations, either with or without adjuvants such as Freund's complete adjuvant and Freund's incomplete adjuvant, and may include naked DNA immunization. Moss, *Semin. Immunol.* 2: 317-327 (1990).

Antibodies from non-human mammals and avian species can be polyclonal or monoclonal, with polyclonal antibodies having certain advantages in immunohistochemical detection of the polypeptides of the present invention and monoclonal antibodies having advantages in identifying and distinguishing particular epitopes of the polypeptides of the present invention. Antibodies from avian species may have particular advantage in detection of the polypeptides of the present invention, in human serum or tissues. Vikinge *et al.*, *Biosens. Bioelectron.* 13: 1257-1262 (1998). Following immunization, the antibodies of the present invention can be obtained using any art-accepted technique. Such techniques are well known in the art and are described in detail in references such as Coligan, *supra*; Zola, *supra*; Howard *et al.* (eds.), Basic Methods in Antibody Production and Characterization, CRC Press (2000); Harlow, *supra*; Davis (ed.), Monoclonal Antibody Protocols, Vol. 45, Humana Press (1995); Delves (ed.), Antibody Production: Essential Techniques, John Wiley & Son Ltd (1997); and Kenney, Antibody Solution: An Antibody Methods Manual, Chapman & Hall (1997).

Briefly, such techniques include, *inter alia*, production of monoclonal antibodies by hybridomas and expression of antibodies or fragments or derivatives thereof from host cells engineered to express immunoglobulin genes or fragments thereof. These two

methods of production are not mutually exclusive: genes encoding antibodies specific for the polypeptides of the present invention can be cloned from hybridomas and thereafter expressed in other host cells. Nor need the two necessarily be performed together: *e.g.*, genes encoding antibodies specific for the polypeptides of the present invention can be
5 cloned directly from B cells known to be specific for the desired protein, as further described in U.S. Patent No. 5,627,052, the disclosure of which is incorporated herein by reference in its entirety, or from antibody-displaying phage.

Recombinant expression in host cells is particularly useful when fragments or derivatives of the antibodies of the present invention are desired.

10 Host cells for recombinant antibody production of whole antibodies, antibody fragments, or antibody derivatives can be prokaryotic or eukaryotic.

Prokaryotic hosts are particularly useful for producing phage displayed antibodies of the present invention.

The technology of phage-displayed antibodies, in which antibody variable region
15 fragments are fused, for example, to the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13, is by now well-established. *See, e.g.*, Sidhu, *Curr. Opin. Biotechnol.* 11(6): 610-6 (2000); Griffiths *et al.*, *Curr. Opin. Biotechnol.* 9(1): 102-8 (1998); Hoogenboom *et al.*, *Immunotechnology*, 4(1): 1-20 (1998); Rader *et al.*, *Current Opinion in Biotechnology* 8: 503-508 (1997); Aujame *et al.*, *Human*
20 *Antibodies* 8: 155-168 (1997); Hoogenboom, *Trends in Biotechnol.* 15: 62-70 (1997); de Kruif *et al.*, 17: 453-455 (1996); Barbas *et al.*, *Trends in Biotechnol.* 14: 230-234 (1996); Winter *et al.*, *Ann. Rev. Immunol.* 433-455 (1994). Techniques and protocols required to generate, propagate, screen (pan), and use the antibody fragments from such libraries have recently been compiled. *See, e.g.*, Barbas (2001), *supra*; Kay, *supra*; and Abelson, *supra*.

25 Typically, phage-displayed antibody fragments are scFv fragments or Fab fragments; when desired, full length antibodies can be produced by cloning the variable regions from the displaying phage into a complete antibody and expressing the full length antibody in a further prokaryotic or a eukaryotic host cell. Eukaryotic cells are also useful for expression of the antibodies, antibody fragments, and antibody derivatives of the
30 present invention. For example, antibody fragments of the present invention can be produced in *Pichia pastoris* and in *Saccharomyces cerevisiae*. *See, e.g.*, Takahashi *et al.*, *Biosci. Biotechnol. Biochem.* 64(10): 2138-44 (2000); Freyre *et al.*, *J. Biotechnol.* 76(2-3):1 57-63 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2): 117-20

(1999); Pennell *et al.*, *Res. Immunol.* 149(6): 599-603 (1998); Eldin *et al.*, *J. Immunol. Methods.* 201(1): 67-75 (1997);, Frenken *et al.*, *Res. Immunol.* 149(6): 589-99 (1998); and Shusta *et al.*, *Nature Biotechnol.* 16(8): 773-7 (1998).

Antibodies, including antibody fragments and derivatives, of the present invention
5 can also be produced in insect cells. *See, e.g.*, Li *et al.*, *Protein Expr. Purif.* 21(1): 121-8 (2001); Ailor *et al.*, *Biotechnol. Bioeng.* 58(2-3): 196-203 (1998); Hsu *et al.*, *Biotechnol. Prog.* 13(1): 96-104 (1997); Edelman *et al.*, *Immunology* 91(1): 13-9 (1997); and Nesbit *et al.*, *J. Immunol. Methods* 151(1-2): 201-8 (1992).

Antibodies and fragments and derivatives thereof of the present invention can also
10 be produced in plant cells, particularly maize or tobacco, Giddings *et al.*, *Nature Biotechnol.* 18(11): 1151-5 (2000); Gavilondo *et al.*, *Biotechniques* 29(1): 128-38 (2000); Fischer *et al.*, *J. Biol. Regul. Homeost. Agents* 14(2): 83-92 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2): 113-6 (1999); Fischer *et al.*, *Biol. Chem.* 380(7-8): 825-39 (1999); Russell, *Curr. Top. Microbiol. Immunol.* 240: 119-38 (1999); and Ma *et al.*, *Plant Physiol.* 109(2): 341-6 (1995).

Antibodies, including antibody fragments and derivatives, of the present invention can also be produced in transgenic, non-human, mammalian milk. *See, e.g.* Pollock *et al.*, *J. Immunol Methods.* 231: 147-57 (1999); Young *et al.*, *Res. Immunol.* 149: 609-10 (1998); and Limonta *et al.*, *Immunotechnology* 1: 107-13 (1995).

20 Mammalian cells useful for recombinant expression of antibodies, antibody fragments, and antibody derivatives of the present invention include CHO cells, COS cells, 293 cells, and myeloma cells. Verma *et al.*, *J. Immunol. Methods* 216(1-2):165-81 (1998) review and compare bacterial, yeast, insect and mammalian expression systems for expression of antibodies. Antibodies of the present invention can also be prepared by cell
25 free translation, as further described in Merk *et al.*, *J. Biochem. (Tokyo)* 125(2): 328-33 (1999) and Ryabova *et al.*, *Nature Biotechnol.* 15(1): 79-84 (1997), and in the milk of transgenic animals, as further described in Pollock *et al.*, *J. Immunol. Methods* 231(1-2): 147-57 (1999).

The invention further provides antibody fragments that bind specifically to one or
30 more of the polypeptides of the present invention or to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid

molecules of the present invention. Among such useful fragments are Fab, Fab', Fv, F(ab)'₂, and single chain Fv (scFv) fragments. Other useful fragments are described in Hudson, *Curr. Opin. Biotechnol.* 9(4): 395-402 (1998).

5 The present invention also relates to antibody derivatives that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention.

10 Among such useful derivatives are chimeric, primatized, and humanized antibodies; such derivatives are less immunogenic in human beings, and thus are more suitable for *in vivo* administration, than are unmodified antibodies from non-human mammalian species. Another useful method is PEGylation to increase the serum half life of the antibodies.

15 Chimeric antibodies typically include heavy and/or light chain variable regions (including both CDR and framework residues) of immunoglobulins of one species, typically mouse, fused to constant regions of another species, typically human. *See, e.g., Morrison et al., Proc. Natl. Acad. Sci USA* 81(21): 6851-5 (1984); Sharon *et al., Nature* 309(5966): 364-7 (1984); Takeda *et al., Nature* 314(6010): 452-4 (1985); and U.S. Patent
20 No. 5,807,715 the disclosure of which is incorporated herein by reference in its entirety. Primatized and humanized antibodies typically include heavy and/or light chain CDRs from a murine antibody grafted into a non-human primate or human antibody V region framework, usually further comprising a human constant region, Riechmann *et al., Nature* 332(6162): 323-7 (1988); Co *et al., Nature* 351(6326): 501-2 (1991); and U.S. Patent Nos.
25 6,054,297; 5,821,337; 5,770,196; 5,766,886; 5,821,123; 5,869,619; 6,180,377; 6,013,256; 5,693,761; and 6,180,370, the disclosures of which are incorporated herein by reference in their entireties. Other useful antibody derivatives of the invention include heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies), single-chain diabodies, and intrabodies.

30 It is contemplated that the nucleic acids encoding the antibodies of the present invention can be operably joined to other nucleic acids forming a recombinant vector for cloning or for expression of the antibodies of the invention. Accordingly, the present invention includes any recombinant vector containing the coding sequences, or part

thereof, whether for eukaryotic transduction, transfection or gene therapy. Such vectors may be prepared using conventional molecular biology techniques, known to those with skill in the art, and would comprise DNA encoding sequences for the immunoglobulin V-regions including framework and CDRs or parts thereof, and a suitable promoter either
5 with or without a signal sequence for intracellular transport. Such vectors may be transduced or transfected into eukaryotic cells or used for gene therapy (Marasco et al., *Proc. Natl. Acad. Sci. (USA)* 90: 7889-7893 (1993); Duan et al., *Proc. Natl. Acad. Sci. (USA)* 91: 5075-5079 (1994), by conventional techniques, known to those with skill in the art.

10 The antibodies of the present invention, including fragments and derivatives thereof, can usefully be labeled. It is, therefore, another aspect of the present invention to provide labeled antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited
15 by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention. The choice of label depends, in part, upon the desired use.

For example, when the antibodies of the present invention are used for immunohistochemical staining of tissue samples, the label can usefully be an enzyme that
20 catalyzes production and local deposition of a detectable product. Enzymes typically conjugated to antibodies to permit their immunohistochemical visualization are well known, and include alkaline phosphatase, β -galactosidase, glucose oxidase, horseradish peroxidase (HRP), and urease. Typical substrates for production and deposition of visually detectable products include o-nitrophenyl-beta-D-galactopyranoside (ONPG);
25 o-phenylenediamine dihydrochloride (OPD); p-nitrophenyl phosphate (PNPP); p-nitrophenyl-beta-D-galactopyranoside (PNPG); 3',3'-diaminobenzidine (DAB); 3-amino-9-ethylcarbazole (AEC); 4-chloro-1-naphthol (CN);
5-bromo-4-chloro-3-indolyl-phosphate (BCIP); ABTS®; BluoGal; iodonitrotetrazolium (INT); nitroblue tetrazolium chloride (NBT); phenazine methosulfate (PMS);
30 phenolphthalein monophosphate (PMP); tetramethyl benzidine (TMB); tetranitroblue tetrazolium (TNBT); X-Gal; X-Gluc; and X-Glucoside.

Other substrates can be used to produce products for local deposition that are luminescent. For example, in the presence of hydrogen peroxide (H₂O₂), horseradish

peroxidase (HRP) can catalyze the oxidation of cyclic diacylhydrazides, such as luminol. Immediately following the oxidation, the luminol is in an excited state (intermediate reaction product), which decays to the ground state by emitting light. Strong enhancement of the light emission is produced by enhancers, such as phenolic compounds. Advantages
5 include high sensitivity, high resolution, and rapid detection without radioactivity and requiring only small amounts of antibody. *See, e.g., Thorpe et al., Methods Enzymol.* 133: 331-53 (1986); Kricka *et al., J. Immunoassay* 17(1): 67-83 (1996); and Lundqvist *et al., J. Biolumin. Chemilumin.* 10(6): 353-9 (1995). Kits for such enhanced chemiluminescent detection (ECL) are available commercially. The antibodies can also be labeled using
10 colloidal gold.

As another example, when the antibodies of the present invention are used, *e.g.,* for flow cytometric detection, for scanning laser cytometric detection, or for fluorescent immunoassay, they can usefully be labeled with fluorophores. There are a wide variety of fluorophore labels that can usefully be attached to the antibodies of the present invention.
15 For flow cytometric applications, both for extracellular detection and for intracellular detection, common useful fluorophores can be fluorescein isothiocyanate (FITC), allophycocyanin (APC), R-phycoerythrin (PE), peridinin chlorophyll protein (PerCP), Texas Red, Cy3, Cy5, fluorescence resonance energy tandem fluorophores such as PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-Texas Red, and APC-Cy7.

Other fluorophores include, *inter alia*, Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY
25 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, OR, USA), and Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, all of
30 which are also useful for fluorescently labeling the antibodies of the present invention. For secondary detection using labeled avidin, streptavidin, captavidin or neutravidin, the antibodies of the present invention can usefully be labeled with biotin.

When the antibodies of the present invention are used, *e.g.*, for western blotting applications, they can usefully be labeled with radioisotopes, such as ^{33}P , ^{32}P , ^{35}S , ^3H , and ^{125}I . As another example, when the antibodies of the present invention are used for radioimmunotherapy, the label can usefully be ^{228}Th , ^{227}Ac , ^{225}Ac , ^{223}Ra , ^{213}Bi , ^{212}Pb , ^{212}Bi , ^{211}At , ^{203}Pb , ^{194}Os , ^{188}Re , ^{186}Re , ^{153}Sm , ^{149}Tb , ^{131}I , ^{125}I , ^{111}In , ^{105}Rh , $^{99\text{m}}\text{Tc}$, ^{97}Ru , ^{90}Y , ^{90}Sr , ^{88}Y , ^{72}Se , ^{67}Cu , or ^{47}Sc .

As another example, when the antibodies of the present invention are to be used for *in vivo* diagnostic use, they can be rendered detectable by conjugation to MRI contrast agents, such as gadolinium diethylenetriaminepentaacetic acid (DTPA), Lauffer *et al.*, *Radiology* 207(2): 529-38 (1998), or by radioisotopic labeling.

As would be understood, use of the labels described above is not restricted to the application as for which they were mentioned.

The antibodies of the present invention, including fragments and derivatives thereof, can also be conjugated to toxins, in order to target the toxin's ablative action to cells that display and/or express the polypeptides of the present invention. Commonly, the antibody in such immunotoxins is conjugated to *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin. See Hall (ed.), Immunotoxin Methods and Protocols (Methods in Molecular Biology, vol. 166), Humana Press (2000); and Frankel *et al.* (eds.), Clinical Applications of Immunotoxins, Springer-Verlag (1998).

The antibodies of the present invention can usefully be attached to a substrate, and it is, therefore, another aspect of the invention to provide antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, attached to a substrate. Substrates can be porous or nonporous, planar or nonplanar. For example, the antibodies of the present invention can usefully be conjugated to filtration media, such as NHS-activated Sepharose or CNBr-activated Sepharose for purposes of immunoaffinity chromatography. For example, the antibodies of the present invention can usefully be attached to paramagnetic microspheres, typically by biotin-streptavidin interaction, which microsphere can then be used for isolation of cells that express or display the polypeptides of the present invention. As

another example, the antibodies of the present invention can usefully be attached to the surface of a microtiter plate for ELISA.

As noted above, the antibodies of the present invention can be produced in prokaryotic and eukaryotic cells. It is, therefore, another aspect of the present invention to
5 provide cells that express the antibodies of the present invention, including hybridoma cells, B cells, plasma cells, and host cells recombinantly modified to express the antibodies of the present invention.

In yet a further aspect, the present invention provides aptamers evolved to bind specifically to one or more of the CSPs of the present invention or to polypeptides
10 encoded by the CSNAs of the invention.

In sum, one of skill in the art, provided with the teachings of this invention, has available a variety of methods which may be used to alter the biological properties of the antibodies of this invention including methods which would increase or decrease the stability or half-life, immunogenicity, toxicity, affinity or yield of a given antibody
15 molecule, or to alter it in any other way that may render it more suitable for a particular application.

Transgenic Animals and Cells

In another aspect, the invention provides transgenic cells and non-human organisms comprising nucleic acid molecules of the invention. In a preferred
20 embodiment, the transgenic cells and non-human organisms comprise a nucleic acid molecule encoding a CSP. In a preferred embodiment, the CSP comprises an amino acid sequence selected from SEQ ID NO: 113-259, or a fragment, mutein, homologous protein or allelic variant thereof. In another preferred embodiment, the transgenic cells and non-human organism comprise a CSNA of the invention, preferably a CSNA comprising a
25 nucleotide sequence selected from the group consisting of SEQ ID NO: 1-112, or a part, substantially similar nucleic acid molecule, allelic variant or hybridizing nucleic acid molecule thereof.

In another embodiment, the transgenic cells and non-human organisms have a targeted disruption or replacement of the endogenous orthologue of the human CSG. The
30 transgenic cells can be embryonic stem cells or somatic cells. The transgenic non-human organisms can be chimeric, nonchimeric heterozygotes, and nonchimeric homozygotes. Methods of producing transgenic animals are well known in the art. *See, e.g., Hogan et*

al., Manipulating the Mouse Embryo: A Laboratory Manual, 2d ed., Cold Spring Harbor Press (1999); Jackson *et al.*, Mouse Genetics and Transgenics: A Practical Approach, Oxford University Press (2000); and Pinkert, Transgenic Animal Technology: A Laboratory Handbook, Academic Press (1999).

- 5 Any technique known in the art may be used to introduce a nucleic acid molecule of the invention into an animal to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection. (*see, e.g.*, Paterson *et al.*, *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver *et al.*, *Biotechnology* 11: 1263-1270 (1993); Wright *et al.*, *Biotechnology* 9: 830-834 (1991); and U.S. Patent No. 4,873,191, herein incorporated by reference in its entirety); retrovirus-mediated gene transfer into germ lines, blastocysts or embryos (*see, e.g.*, Van der Putten *et al.*, *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)); gene targeting in embryonic stem cells (*see, e.g.*, Thompson *et al.*, *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (*see, e.g.*, Lo, 1983, *Mol. Cell. Biol.* 3: 1803-1814 (1983)); introduction using a gene gun (*see, e.g.*, Ulmer *et al.*, *Science* 259: 1745-49 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (*see, e.g.*, Lavitrano *et al.*, *Cell* 57: 717-723 (1989)).

- 20 Other techniques include, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (*see, e.g.*, Campell *et al.*, *Nature* 380: 64-66 (1996); Wilmut *et al.*, *Nature* 385: 810-813 (1997)). The present invention provides for transgenic animals that carry the transgene (*i.e.*, a nucleic acid molecule of the invention) in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.* e., mosaic animals or chimeric animals.

- 25 The transgene may be integrated as a single transgene or as multiple copies, such as in concatamers, *e. g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, *e.g.*, the teaching of Lasko *et al. et al.*, *Proc. Natl. Acad. Sci. USA* 89: 6232- 6236 (1992). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

- 30 Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the

transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (RT-PCR). Samples of transgenic gene-expressing tissue may also be evaluated

5 immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than

10 one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA

15 analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of

20 the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Methods for creating a transgenic animal with a disruption of a targeted gene are also well known in the art. In general, a vector is designed to comprise some nucleotide

25 sequences homologous to the endogenous targeted gene. The vector is introduced into a cell so that it may integrate, via homologous recombination with chromosomal sequences, into the endogenous gene, thereby disrupting the function of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type. *See, e.g., Gu et al., Science* 265: 103-106

30 (1994). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. *See, e.g., Smithies et al., Nature* 317: 230-234 (1985); Thomas et al., *Cell* 51: 503-512 (1987); Thompson et al., *Cell* 5: 313-321 (1989).

In one embodiment, a mutant, non-functional nucleic acid molecule of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous nucleic acid sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable
5 marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications
10 to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene. *See, e.g.,* Thomas, *supra* and Thompson, *supra*. However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

15 In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient in vivo. Such cells may be obtained from an animal or patient or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells,
20 blood cells (*e.g.*, lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, *e.g.*, by transduction (using viral vectors, and preferably
25 vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve
30 expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, *e.g.*, in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, *e.g.*, genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. *See, e.g.*, U.S. Patent Nos. 5,399,349 and 5,460,959, each of which is
5 incorporated by reference herein in its entirety.

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of
10 components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with
15 aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Computer Readable Means

A further aspect of the invention is a computer readable means for storing the nucleic acid and amino acid sequences of the instant invention. In a preferred
20 embodiment, the invention provides a computer readable means for storing SEQ ID NO: 113-259 and SEQ ID NO: 1-112 as described herein, as the complete set of sequences or in any combination. The records of the computer readable means can be accessed for reading and display and for interface with a computer system for the application of programs allowing for the location of data upon a query for data meeting certain criteria,
25 the comparison of sequences, the alignment or ordering of sequences meeting a set of criteria, and the like.

The nucleic acid and amino acid sequences of the invention are particularly useful as components in databases useful for search analyses as well as in sequence analysis algorithms. As used herein, the terms "nucleic acid sequences of the invention" and
30 "amino acid sequences of the invention" mean any detectable chemical or physical characteristic of a polynucleotide or polypeptide of the invention that is or may be reduced to or stored in a computer readable form. These include, without limitation,

chromatographic scan data or peak data, photographic data or scan data therefrom, and mass spectrographic data.

This invention provides computer readable media having stored thereon sequences of the invention. A computer readable medium may comprise one or more of the
5 following: a nucleic acid sequence comprising a sequence of a nucleic acid sequence of the invention; an amino acid sequence comprising an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence
10 of the invention; a data set representing a nucleic acid sequence comprising the sequence of one or more nucleic acid sequences of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of
15 amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence of the invention; a data set representing a nucleic acid sequence comprising the sequence of a nucleic acid sequence of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention. The computer readable medium can
20 be any composition of matter used to store information or data, including, for example, commercially available floppy disks, tapes, hard drives, compact disks, and video disks.

Also provided by the invention are methods for the analysis of character sequences, particularly genetic sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and
25 similarity analysis, RNA structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, and sequencing chromatogram peak analysis.

A computer-based method is provided for performing nucleic acid sequence identity or similarity identification. This method comprises the steps of providing a
30 nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and comparing said nucleic acid sequence to at least one nucleic acid or amino acid sequence to identify sequence identity or similarity.

A computer-based method is also provided for performing amino acid homology identification, said method comprising the steps of: providing an amino acid sequence comprising the sequence of an amino acid of the invention in a computer readable medium; and comparing said amino acid sequence to at least one nucleic acid or an amino acid sequence to identify homology.

A computer-based method is still further provided for assembly of overlapping nucleic acid sequences into a single nucleic acid sequence, said method comprising the steps of: providing a first nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and screening for at least one overlapping region between said first nucleic acid sequence and a second nucleic acid sequence. In addition, the invention includes a method of using patterns of expression associated with either the nucleic acids or proteins in a computer-based method to diagnose disease.

Diagnostic Methods for Colon Cancer

The present invention also relates to quantitative and qualitative diagnostic assays and methods for detecting, diagnosing, monitoring, staging and predicting cancers by comparing expression of a CSNA or a CSP in a human patient that has or may have colon cancer, or who is at risk of developing colon cancer, with the expression of a CSNA or a CSP in a normal human control. For purposes of the present invention, "expression of a CSNA" or "CSNA expression" means the quantity of CSNA mRNA that can be measured by any method known in the art or the level of transcription that can be measured by any method known in the art in a cell, tissue, organ or whole patient. Similarly, the term "expression of a CSP" or "CSP expression" means the amount of CSP that can be measured by any method known in the art or the level of translation of a CSNA that can be measured by any method known in the art.

The present invention provides methods for diagnosing colon cancer in a patient, in particular adenocarcinoma, by analyzing for changes in levels of CSNA or CSP in cells, tissues, organs or bodily fluids compared with levels of CSNA or CSP in cells, tissues, organs or bodily fluids of preferably the same type from a normal human control, wherein an increase, or decrease in certain cases, in levels of a CSNA or CSP in the patient versus the normal human control is associated with the presence of colon cancer or with a predilection to the disease. In another preferred embodiment, the present invention

provides methods for diagnosing colon cancer in a patient by analyzing changes in the structure of the mRNA of a CSG compared to the mRNA from a normal control. These changes include, without limitation, aberrant splicing, alterations in polyadenylation and/or alterations in 5' nucleotide capping. In yet another preferred embodiment, the present invention provides methods for diagnosing colon cancer in a patient by analyzing changes in a CSP compared to a CSP from a normal patient. These changes include, *e.g.*, alterations, including post translational modifications such as glycosylation and/or phosphorylation of the CSP or changes in the subcellular CSP localization.

For purposes of the present invention, diagnosing means that CSNA or CSP levels are used to determine the presence or absence of disease in a patient. As will be understood by those of skill in the art, measurement of other diagnostic parameters may be required for definitive diagnosis or determination of the appropriate treatment for the disease. The determination may be made by a clinician, a doctor, a testing laboratory, or a patient using an over the counter test. The patient may have symptoms of disease or may be asymptomatic. In addition, the CSNA or CSP levels of the present invention may be used as screening marker to determine whether further tests or biopsies are warranted. In addition, the CSNA or CSP levels may be used to determine the vulnerability or susceptibility to disease.

In a preferred embodiment, the expression of a CSNA is measured by determining the amount of a mRNA that encodes an amino acid sequence selected from SEQ ID NO: 113-259, a homolog, an allelic variant, or a fragment thereof. In a more preferred embodiment, the CSNA expression that is measured is the level of expression of a CSNA mRNA selected from SEQ ID NO: 1-112, or a hybridizing nucleic acid, homologous nucleic acid or allelic variant thereof, or a part of any of these nucleic acid molecules. CSNA expression may be measured by any method known in the art, such as those described *supra*, including measuring mRNA expression by Northern blot, quantitative or qualitative reverse transcriptase PCR (RT-PCR), microarray, dot or slot blots or *in situ* hybridization. *See, e.g.*, Ausubel (1992), *supra*; Ausubel (1999), *supra*; Sambrook (1989), *supra*; and Sambrook (2001), *supra*. CSNA transcription may be measured by any method known in the art including using a reporter gene hooked up to the promoter of a CSG of interest or doing nuclear run-off assays. Alterations in mRNA structure, *e.g.*, aberrant splicing variants, may be determined by any method known in the art, including, RT-PCR followed by sequencing or restriction analysis. As necessary, CSNA expression

may be compared to a known control, such as normal colon nucleic acid, to detect a change in expression.

In another preferred embodiment, the expression of a CSP is measured by determining the level of a CSP having an amino acid sequence selected from the group consisting of SEQ ID NO: 113-259, a homolog, an allelic variant, or a fragment thereof. Such levels are preferably determined in at least one of cells, tissues, organs and/or bodily fluids, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for diagnosing over- or underexpression of a CSNA or CSP compared to normal control bodily fluids, cells, or tissue samples may be used to diagnose the presence of colon cancer. The expression level of a CSP may be determined by any method known in the art, such as those described *supra*. In a preferred embodiment, the CSP expression level may be determined by radioimmunoassays, competitive-binding assays, ELISA, Western blot, FACS, immunohistochemistry, immunoprecipitation, proteomic approaches: two-dimensional gel electrophoresis (2D electrophoresis) and non-gel-based approaches such as mass spectrometry or protein interaction profiling. See, e.g., Harlow (1999), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Alterations in the CSP structure may be determined by any method known in the art, including, e.g., using antibodies that specifically recognize phosphoserine, phosphothreonine or phosphotyrosine residues, two-dimensional polyacrylamide gel electrophoresis (2D PAGE) and/or chemical analysis of amino acid residues of the protein. *Id.*

In a preferred embodiment, a radioimmunoassay (RIA) or an ELISA is used. An antibody specific to a CSP is prepared if one is not already available. In a preferred embodiment, the antibody is a monoclonal antibody. The anti-CSP antibody is bound to a solid support and any free protein binding sites on the solid support are blocked with a protein such as bovine serum albumin. A sample of interest is incubated with the antibody on the solid support under conditions in which the CSP will bind to the anti-CSP antibody. The sample is removed, the solid support is washed to remove unbound material, and an anti-CSP antibody that is linked to a detectable reagent (a radioactive substance for RIA and an enzyme for ELISA) is added to the solid support and incubated under conditions in which binding of the CSP to the labeled antibody will occur. After binding, the unbound labeled antibody is removed by washing. For an ELISA, one or more substrates are added to produce a colored reaction product that is based upon the amount of a CSP in the

sample. For an RIA, the solid support is counted for radioactive decay signals by any method known in the art. Quantitative results for both RIA and ELISA typically are obtained by reference to a standard curve.

Other methods to measure CSP levels are known in the art. For instance, a
5 competition assay may be employed wherein an anti-CSP antibody is attached to a solid support and an allocated amount of a labeled CSP and a sample of interest are incubated with the solid support. The amount of labeled CSP attached to the solid support can be correlated to the quantity of a CSP in the sample.

Of the proteomic approaches, 2D PAGE is a well known technique. Isolation of
10 individual proteins from a sample such as serum is accomplished using sequential separation of proteins by isoelectric point and molecular weight. Typically, polypeptides are first separated by isoelectric point (the first dimension) and then separated by size using an electric current (the second dimension). In general, the second dimension is perpendicular to the first dimension. Because no two proteins with different sequences are
15 identical on the basis of both size and charge, the result of 2D PAGE is a roughly square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent protein microsequencing can reveal the relative abundance of a given protein and the identity of the proteins in the sample.

Expression levels of a CSNA can be determined by any method known in the art,
20 including PCR and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of thousands of other mRNA species. In
25 RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA) with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction.

Hybridization to specific DNA molecules (*e.g.*, oligonucleotides) arrayed on a solid support can be used to both detect the expression of and quantitate the level of
30 expression of one or more CSNAs of interest. In this approach, all or a portion of one or more CSNAs is fixed to a substrate. A sample of interest, which may comprise RNA, *e.g.*, total RNA or polyA-selected mRNA, or a complementary DNA (cDNA) copy of the RNA is incubated with the solid support under conditions in which hybridization will occur

between the DNA on the solid support and the nucleic acid molecules in the sample of interest. Hybridization between the substrate-bound DNA and the nucleic acid molecules in the sample can be detected and quantitated by several means, including, without limitation, radioactive labeling or fluorescent labeling of the nucleic acid molecule or a
5 secondary molecule designed to detect the hybrid.

The above tests can be carried out on samples derived from a variety of cells, bodily fluids and/or tissue extracts such as homogenates or solubilized tissue obtained from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any
10 other bodily secretion or derivative thereof. As used herein "blood" includes whole blood, plasma, serum, circulating epithelial cells, constituents, or any derivative of blood.

In addition to detection in bodily fluids, the proteins and nucleic acids of the invention are suitable to detection by cell capture technology. Whole cells may be captured by a variety methods for example magnetic separation, such as described in U.S.
15 Patent. Nos. 5,200,084; 5,186,827; 5,108,933; and 4,925,788, the disclosures of which are incorporated herein by reference in their entirety. Epithelial cells may be captured using such products as Dynabeads® or CELlection™ (Dynal Biotech, Oslo, Norway). Alternatively, fractions of blood may be captured, e.g., the buffy coat fraction (50mm cells isolated from 5ml of blood) containing epithelial cells. In addition, cancer cells may be
20 captured using the techniques described in WO 00/47998, the disclosure of which is incorporated herein by reference in its entirety. Once the cells are captured or concentrated, the proteins or nucleic acids are detected by the means described in the subject application. Alternatively, nucleic acids may be captured directly from blood samples, see U.S. Patent Nos. 6,156,504, 5,501,963; or WO 01/42504 , the disclosures of
25 which are incorporated herein by reference in their entirety.

In a preferred embodiment, the specimen tested for expression of CSNA or CSP includes without limitation colon tissue, fecal samples, colonocytes, colon cells grown in cell culture, blood, serum, lymph node tissue, and lymphatic fluid. In another preferred embodiment, especially when metastasis of a primary colon cancer is known or suspected,
30 specimens include, without limitation, tissues from brain, bone, bone marrow, liver, lungs, and adrenal glands. In general, the tissues may be sampled by biopsy, including, without limitation, needle biopsy, e.g., transthoracic needle aspiration, cervical mediastinoscopy,

endoscopic lymph node biopsy, video-assisted thoracoscopy, exploratory thoracotomy, bone marrow biopsy and bone marrow aspiration.

Colonocytes represent an important source of the CSP or CSNA because they provide a picture of the immediate past metabolic history of the GI tract of a subject. In addition, such cells are representative of the cell population from a statistically large sampling frame reflecting the state of the colonic mucosa along the entire length of the colon in a non-invasive manner, in contrast to a limited sampling by colonic biopsy using an invasive procedure involving endoscopy. Specific examples of patents describing the isolation of colonocytes include U.S. Patent Nos. 6,335,193; 6,020,137 5,741,650; 6,258,541; US 2001 0026925 A1; WO 00/63358 A1, the disclosures of which are incorporated herein by reference in their entireties.

All the methods of the present invention may optionally include determining the expression levels of one or more other cancer markers in addition to determining the expression level of a CSNAs or CSP. In many cases, the use of another cancer marker will decrease the likelihood of false positives or false negatives. In one embodiment, the one or more other cancer markers include other CSNA or CSPs as disclosed herein. Other cancer markers useful in the present invention will depend on the cancer being tested and are known to those of skill in the art. In a preferred embodiment, at least one other cancer marker in addition to a particular CSNA or CSP is measured. In a more preferred embodiment, at least two other additional cancer markers are used. In an even more preferred embodiment, at least three, more preferably at least five, even more preferably at least ten additional cancer markers are used.

Diagnosing

In one aspect, the invention provides a method for determining the expression levels and/or structural alterations of one or more CSNA and/or CSP in a sample from a patient suspected of having colon cancer. In general, the method comprises the steps of obtaining the sample from the patient, determining the expression level or structural alterations of a CSNA and/or CSP and then ascertaining whether the patient has colon cancer from the expression level of the CSNA or CSP. In general, if high expression relative to a control of a CSNA or CSP is indicative of colon cancer, a diagnostic assay is considered positive if the level of expression of the CSNA or CSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five

times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a CSNA or CSP is indicative of colon cancer, a diagnostic assay is considered positive if the level of expression of the CSNA or CSP is at least one and a half
5 times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

The present invention also provides a method of determining whether colon cancer
10 has metastasized in a patient. One may identify whether the colon cancer has metastasized by measuring the expression levels and/or structural alterations of one or more CSNAs and/or CSPs in a variety of tissues. The presence of a CSNA or CSP in a tissue other than colon at levels higher than that of corresponding noncancerous tissue (*e.g.*, the same tissue from another individual) is indicative of metastasis if high level expression of a CSNA or
15 CSP is associated with colon cancer. Similarly, the presence of a CSNA or CSP in a tissue other than colon at levels lower than that of corresponding noncancerous tissue is indicative of metastasis if low level expression of a CSNA or CSP is associated with colon cancer. Further, the presence of a structurally altered CSNA or CSP that is associated with colon cancer is also indicative of metastasis.

20 In general, if high expression relative to a control of a CSNA or CSP is indicative of metastasis, an assay for metastasis is considered positive if the level of expression of the CSNA or CSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human
25 control. In contrast, if low expression relative to a control of a CSNA or CSP is indicative of metastasis, an assay for metastasis is considered positive if the level of expression of the CSNA or CSP is at least one and a half times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human
30 control.

Staging

The invention also provides a method of staging colon cancer in a human patient. The method comprises identifying a human patient having colon cancer and analyzing cells, tissues or bodily fluids from such human patient for expression levels and/or structural alterations of one or more CSNAs or CSPs. First, one or more tumors from a variety of patients are staged according to procedures well known in the art, and the expression levels of one or more CSNAs or CSPs is determined for each stage to obtain a standard expression level for each CSNA and CSP. Then, the CSNA or CSP expression levels of the CSNA or CSP are determined in a biological sample from a patient whose stage of cancer is not known. The CSNA or CSP expression levels from the patient are then compared to the standard expression level. By comparing the expression level of the CSNAs and CSPs from the patient to the standard expression levels, one may determine the stage of the tumor. The same procedure may be followed using structural alterations of a CSNA or CSP to determine the stage of a colon cancer.

Monitoring

Further provided is a method of monitoring colon cancer in a human patient. One may monitor a human patient to determine whether there has been metastasis and, if there has been, when metastasis began to occur. One may also monitor a human patient to determine whether a preneoplastic lesion has become cancerous. One may also monitor a human patient to determine whether a therapy, *e.g.*, chemotherapy, radiotherapy or surgery, has decreased or eliminated the colon cancer. The monitoring may determine if there has been a reoccurrence and, if so, determine its nature. The method comprises identifying a human patient that one wants to monitor for colon cancer, periodically analyzing cells, tissues or bodily fluids from such human patient for expression levels of one or more CSNAs or CSPs, and comparing the CSNA or CSP levels over time to those CSNA or CSP expression levels obtained previously. Patients may also be monitored by measuring one or more structural alterations in a CSNA or CSP that are associated with colon cancer.

If increased expression of a CSNA or CSP is associated with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting an increase in the expression level of a CSNA or CSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. One having ordinary skill in the art would recognize that if this were the case, then a decreased

expression level would be indicative of no metastasis, effective therapy or failure to progress to a neoplastic lesion. If decreased expression of a CSNA or CSP is associated with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting a decrease in the expression level of a CSNA or CSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. In a preferred embodiment, the levels of CSNAs or CSPs are determined from the same cell type, tissue or bodily fluid as prior patient samples. Monitoring a patient for onset of colon cancer metastasis is periodic and preferably is done on a quarterly basis, but may be done more or less frequently.

The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with increased or decreased expression levels of a CSNA and/or CSP. The present invention provides a method in which a test sample is obtained from a human patient and one or more CSNAs and/or CSPs are detected. The presence of higher (or lower) CSNA or CSP levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly colon cancer. The effectiveness of therapeutic agents to decrease (or increase) expression or activity of one or more CSNAs and/or CSPs of the invention can also be monitored by analyzing levels of expression of the CSNAs and/or CSPs in a human patient in clinical trials or in *in vitro* screening assays such as in human cells. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the human patient or cells, as the case may be, to the agent being tested.

Detection of Genetic Lesions or Mutations

The methods of the present invention can also be used to detect genetic lesions or mutations in a CSG, thereby determining if a human with the genetic lesion is susceptible to developing colon cancer or to determine what genetic lesions are responsible, or are partly responsible, for a person's existing colon cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion, insertion and/or substitution of one or more nucleotides from the CSGs of this invention, a chromosomal rearrangement of a CSG, an aberrant modification of a CSG (such as of the methylation pattern of the genomic DNA), or allelic loss of a CSG. Methods to detect such lesions in the CSG of

this invention are known to those having ordinary skill in the art following the teachings of the specification.

Methods of Detecting Noncancerous Colon Diseases

The present invention also provides methods for determining the expression levels
5 and/or structural alterations of one or more CSNAs and/or CSPs in a sample from a patient suspected of having or known to have a noncancerous colon disease. In general, the method comprises the steps of obtaining a sample from the patient, determining the expression level or structural alterations of a CSNA and/or CSP, comparing the expression level or structural alteration of the CSNA or CSP to a normal colon control, and then
10 ascertaining whether the patient has a noncancerous colon disease. In general, if high expression relative to a control of a CSNA or CSP is indicative of a particular noncancerous colon disease, a diagnostic assay is considered positive if the level of expression of the CSNA or CSP is at least two times higher, and more preferably are at least five times higher, even more preferably at least ten times higher, than in preferably
15 the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a CSNA or CSP is indicative of a noncancerous colon disease, a diagnostic assay is considered positive if the level of expression of the CSNA or CSP is at least two times lower, more preferably are at least five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid
20 of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

One having ordinary skill in the art may determine whether a CSNA and/or CSP is associated with a particular noncancerous colon disease by obtaining colon tissue from a patient having a noncancerous colon disease of interest and determining which CSNAs
25 and/or CSPs are expressed in the tissue at either a higher or a lower level than in normal colon tissue. In another embodiment, one may determine whether a CSNA or CSP exhibits structural alterations in a particular noncancerous colon disease state by obtaining colon tissue from a patient having a noncancerous colon disease of interest and determining the structural alterations in one or more CSNAs and/or CSPs relative to
30 normal colon tissue.

Methods for Identifying Colon Tissue

In another aspect, the invention provides methods for identifying colon tissue. These methods are particularly useful in, *e.g.*, forensic science, colon cell differentiation and development, and in tissue engineering.

- 5 In one embodiment, the invention provides a method for determining whether a sample is colon tissue or has colon tissue-like characteristics. The method comprises the steps of providing a sample suspected of comprising colon tissue or having colon tissue-like characteristics, determining whether the sample expresses one or more CSNAs and/or CSPs, and, if the sample expresses one or more CSNAs and/or CSPs, concluding that the
- 10 sample comprises colon tissue. In a preferred embodiment, the CSNA encodes a polypeptide having an amino acid sequence selected from SEQ ID NO: 113-259, or a homolog, allelic variant or fragment thereof. In a more preferred embodiment, the CSNA has a nucleotide sequence selected from SEQ ID NO: 1-112, or a hybridizing nucleic acid, an allelic variant or a part thereof. Determining whether a sample expresses a CSNA can
- 15 be accomplished by any method known in the art. Preferred methods include hybridization to microarrays, Northern blot hybridization, and quantitative or qualitative RT-PCR. In another preferred embodiment, the method can be practiced by determining whether a CSP is expressed. Determining whether a sample expresses a CSP can be accomplished by any method known in the art. Preferred methods include Western blot,
- 20 ELISA, RIA and 2D PAGE. In one embodiment, the CSP has an amino acid sequence selected from SEQ ID NO: 113-259, or a homolog, allelic variant or fragment thereof. In another preferred embodiment, the expression of at least two CSNAs and/or CSPs is determined. In a more preferred embodiment, the expression of at least three, more preferably four and even more preferably five CSNAs and/or CSPs are determined.
- 25 In one embodiment, the method can be used to determine whether an unknown tissue is colon tissue. This is particularly useful in forensic science, in which small, damaged pieces of tissues that are not identifiable by microscopic or other means are recovered from a crime or accident scene. In another embodiment, the method can be used to determine whether a tissue is differentiating or developing into colon tissue. This
- 30 is important in monitoring the effects of the addition of various agents to cell or tissue culture, *e.g.*, in producing new colon tissue by tissue engineering. These agents include, *e.g.*, growth and differentiation factors, extracellular matrix proteins and culture medium. Other factors that may be measured for effects on tissue development and differentiation

include gene transfer into the cells or tissues, alterations in pH, aqueous:air interface and various other culture conditions.

Methods for Producing and Modifying Colon Tissue

In another aspect, the invention provides methods for producing engineered colon
5 tissue or cells. In one embodiment, the method comprises the steps of providing cells, introducing a CSNA or a CSG into the cells, and growing the cells under conditions in which they exhibit one or more properties of colon tissue cells. In a preferred embodiment, the cells are pluripotent. As is well known in the art, normal colon tissue comprises a large number of different cell types. Thus, in one embodiment, the
10 engineered colon tissue or cells comprises one of these cell types. In another embodiment, the engineered colon tissue or cells comprises more than one colon cell type. Further, the culture conditions of the cells or tissue may require manipulation in order to achieve full differentiation and development of the colon cell tissue. Methods for manipulating culture conditions are well known in the art.

15 Nucleic acid molecules encoding one or more CSPs are introduced into cells, preferably pluripotent cells. In a preferred embodiment, the nucleic acid molecules encode CSPs having amino acid sequences selected from SEQ ID NO: 113-259, or homologous proteins, analogs, allelic variants or fragments thereof. In a more preferred embodiment, the nucleic acid molecules have a nucleotide sequence selected from SEQ ID
20 NO: 1-112, or hybridizing nucleic acids, allelic variants or parts thereof. In another highly preferred embodiment, a CSG is introduced into the cells. Expression vectors and methods of introducing nucleic acid molecules into cells are well known in the art and are described in detail, *supra*.

Artificial colon tissue may be used to treat patients who have lost some or all of
25 their colon function.

Pharmaceutical Compositions

In another aspect, the invention provides pharmaceutical compositions comprising the nucleic acid molecules, polypeptides, fusion proteins, antibodies, antibody derivatives, antibody fragments, agonists, antagonists, or inhibitors of the present invention. In a
30 preferred embodiment, the pharmaceutical composition comprises a CSNA or part thereof. In a more preferred embodiment, the CSNA has a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-112, a nucleic acid that hybridizes thereto, an allelic

variant thereof, or a nucleic acid that has substantial sequence identity thereto. In another preferred embodiment, the pharmaceutical composition comprises a CSP or fragment thereof. In a more preferred embodiment, the pharmaceutical composition comprises a CSP having an amino acid sequence that is selected from the group consisting of SEQ ID NO: 113-259, a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof. In another preferred embodiment, the pharmaceutical composition comprises an anti-CSP antibody, preferably an antibody that specifically binds to a CSP having an amino acid that is selected from the group consisting of SEQ ID NO: 113-259, or an antibody that binds to a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof.

Due to the association of angiogenesis with cancer vascularization there is great need of new markers and methods for diagnosing angiogenesis activity to identify developing tumors and angiogenesis related diseases. Furthermore, great need is also present for new molecular targets useful in the treatment of angiogenesis and angiogenesis related diseases such as cancer. In addition known modulators of angiogenesis such as endostatin or vascular endothelial growth factor (VEGF). Use of the methods and compositions disclosed herein in combination with anti-angiogenesis drugs, drugs that block the matrix breakdown (such as BMS-275291, Dalteparin (Fragmin®), Suramin), drugs that inhibit endothelial cells (2-methoxyestradiol (2-ME), CC-5013 (Thalidomide Analog), Combretastatin A4 Phosphate, LY317615 (Protein Kinase C Beta Inhibitor), Soy Isoflavone (Genistein; Soy Protein Isolate), Thalidomide), drugs that block activators of angiogenesis (AE-941 (Neovastat™; GW786034), Anti-VEGF Antibody (Bevacizumab; Avastin™), Interferon-alpha, PTK787/ZK 222584, VEGF-Trap, ZD6474), Drugs that inhibit endothelial-specific integrin/survival signaling (EMD 121974, Anti-Anb3 Integrin Antibody (Medi-522; Vitaxin™)).

Such a composition typically contains from about 0.1 to 90% by weight of a therapeutic agent of the invention formulated in and/or with a pharmaceutically acceptable carrier or excipient.

Pharmaceutical formulation is a well-established art that is further described in Gennaro (ed.), Remington: The Science and Practice of Pharmacy, 20th ed., Lippincott, Williams & Wilkins (2000); Ansel *et al.*, Pharmaceutical Dosage Forms and Drug Delivery Systems, 7th ed., Lippincott Williams & Wilkins (1999); and Kibbe (ed.),

Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3rd ed. (2000) and thus need not be described in detail herein.

Briefly, formulation of the pharmaceutical compositions of the present invention will depend upon the route chosen for administration. The pharmaceutical compositions
5 utilized in this invention can be administered by various routes including both enteral and parenteral routes, including oral, intravenous, intramuscular, subcutaneous, inhalation, topical, sublingual, rectal, intra-arterial, intramedullary, intrathecal, intraventricular, transmucosal, transdermal, intranasal, intraperitoneal, intrapulmonary, and intrauterine.

Oral dosage forms can be formulated as tablets, pills, dragees, capsules, liquids,
10 gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, sodium
15 carboxymethylcellulose, or microcrystalline cellulose; gums including arabic and tragacanth; proteins such as gelatin and collagen; inorganics, such as kaolin, calcium carbonate, dicalcium phosphate, sodium chloride; and other agents such as acacia and alginic acid.

Agents that facilitate disintegration and/or solubilization can be added, such as the
20 cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate, microcrystalline cellulose, cornstarch, sodium starch glycolate, and alginic acid.

Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone (PovidoneTM), hydroxypropyl methylcellulose, sucrose, starch and ethylcellulose.

25 Lubricants that can be used include magnesium stearates, stearic acid, silicone fluid, talc, waxes, oils, and colloidal silica.

Fillers, agents that facilitate disintegration and/or solubilization, tablet binders and lubricants, including the aforementioned, can be used singly or in combination.

Solid oral dosage forms need not be uniform throughout. For example, dragee
30 cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which can also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

Oral dosage forms of the present invention include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Additionally, dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Liquid formulations of the pharmaceutical compositions for oral (enteral) administration are prepared in water or other aqueous vehicles and can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin, carrageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents.

The pharmaceutical compositions of the present invention can also be formulated for parenteral administration. Formulations for parenteral administration can be in the form of aqueous or non-aqueous isotonic sterile injection solutions or suspensions.

For intravenous injection, water soluble versions of the compounds of the present invention are formulated in, or if provided as a lyophilate, mixed with, a physiologically acceptable fluid vehicle, such as 5% dextrose ("D5"), physiologically buffered saline, 0.9% saline, Hanks' solution, or Ringer's solution. Intravenous formulations may include carriers, excipients or stabilizers including, without limitation, calcium, human serum albumin, citrate, acetate, calcium chloride, carbonate, and other salts.

Intramuscular preparations, *e.g.* a sterile formulation of a suitable soluble salt form of the compounds of the present invention, can be dissolved and administered in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. Alternatively, a suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid (*e.g.*, ethyl oleate), fatty oils such as sesame oil, triglycerides, or liposomes.

Parenteral formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate,

isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like).

Aqueous injection suspensions can also contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran.

- 5 Non-lipid polycationic amino polymers can also be used for delivery. Optionally, the suspension can also contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

- 10 Pharmaceutical compositions of the present invention can also be formulated to permit injectable, long-term, deposition. Injectable depot forms may be made by forming microencapsulated matrices of the compound in biodegradable polymers such as polylactide-polyglycolide. Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are also prepared by entrapping the drug in microemulsions that
15 are compatible with body tissues.

- The pharmaceutical compositions of the present invention can be administered topically. For topical use the compounds of the present invention can also be prepared in suitable forms to be applied to the skin, or mucus membranes of the nose and throat, and can take the form of lotions, creams, ointments, liquid sprays or inhalants, drops, tinctures,
20 lozenges, or throat paints. Such topical formulations further can include chemical compounds such as dimethylsulfoxide (DMSO) to facilitate surface penetration of the active ingredient. In other transdermal formulations, typically in patch-delivered formulations, the pharmaceutically active compound is formulated with one or more skin penetrants, such as 2-N-methyl-pyrrolidone (NMP) or Azone. A topical semi-solid
25 ointment formulation typically contains a concentration of the active ingredient from about 1 to 20%, *e.g.*, 5 to 10%, in a carrier such as a pharmaceutical cream base.

For application to the eyes or ears, the compounds of the present invention can be presented in liquid or semi-liquid form formulated in hydrophobic or hydrophilic bases as ointments, creams, lotions, paints or powders.

- 30 For rectal administration the compounds of the present invention can be administered in the form of suppositories admixed with conventional carriers such as cocoa butter, wax or other glyceride.

Inhalation formulations can also readily be formulated. For inhalation, various powder and liquid formulations can be prepared. For aerosol preparations, a sterile formulation of the compound or salt form of the compound may be used in inhalers, such as metered dose inhalers, and nebulizers. Aerosolized forms may be especially useful for treating respiratory disorders.

Alternatively, the compounds of the present invention can be in powder form for reconstitution in the appropriate pharmaceutically acceptable carrier at the time of delivery.

The pharmaceutically active compound in the pharmaceutical compositions of the present invention can be provided as the salt of a variety of acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms.

After pharmaceutical compositions have been prepared, they are packaged in an appropriate container and labeled for treatment of an indicated condition.

The active compound will be present in an amount effective to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

A "therapeutically effective dose" refers to that amount of active ingredient, for example CSP polypeptide, fusion protein, or fragments thereof, antibodies specific for CSP, agonists, antagonists or inhibitors of CSP, which ameliorates the signs or symptoms of the disease or prevent progression thereof; as would be understood in the medical arts, cure, although desired, is not required.

The therapeutically effective dose of the pharmaceutical agents of the present invention can be estimated initially by *in vitro* tests, such as cell culture assays, followed by assay in model animals, usually mice, rats, rabbits, dogs, or pigs. The animal model can also be used to determine an initial preferred concentration range and route of administration.

For example, the ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population) can be determined in one or more cell culture of animal model systems. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as LD50/ED50. Pharmaceutical compositions that exhibit large therapeutic indices are preferred.

The data obtained from cell culture assays and animal studies are used in formulating an initial dosage range for human use, and preferably provide a range of circulating concentrations that includes the ED50 with little or no toxicity. After administration, or between successive administrations, the circulating concentration of active agent varies within this range depending upon pharmacokinetic factors well known in the art, such as the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors specific to the subject requiring treatment. Factors that can be taken into account by the practitioner include the severity of the disease state, general health of the subject, age, weight, gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Where the therapeutic agent is a protein or antibody of the present invention, the therapeutic protein or antibody agent typically is administered at a daily dosage of 0.01 mg to 30 mg/kg of body weight of the patient (*e.g.*, 1mg/kg to 5 mg/kg). The pharmaceutical formulation can be administered in multiple doses per day, if desired, to achieve the total desired daily dose.

Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

Conventional methods, known to those of ordinary skill in the art of medicine, can be used to administer the pharmaceutical formulation(s) of the present invention to the patient. The pharmaceutical compositions of the present invention can be administered alone, or in combination with other therapeutic agents or interventions.

30 Therapeutic Methods

The present invention further provides methods of treating subjects having defects in a gene of the invention, *e.g.*, in expression, activity, distribution, localization, and/or

solubility, which can manifest as a disorder of colon function. As used herein, "treating" includes all medically-acceptable types of therapeutic intervention, including palliation and prophylaxis (prevention) of disease. The term "treating" encompasses any improvement of a disease, including minor improvements. These methods are discussed
5 below.

Gene Therapy and Vaccines

The isolated nucleic acids of the present invention can also be used to drive *in vivo* expression of the polypeptides of the present invention. *In vivo* expression can be driven from a vector, typically a viral vector, often a vector based upon a replication incompetent
10 retrovirus, an adenovirus, or an adeno-associated virus (AAV), for the purpose of gene therapy. *In vivo* expression can also be driven from signals endogenous to the nucleic acid or from a vector, often a plasmid vector, such as pVAX1 (Invitrogen, Carlsbad, CA, USA), for purpose of "naked" nucleic acid vaccination, as further described in U.S. Patent Nos. 5,589,466; 5,679,647; 5,804,566; 5,830,877; 5,843,913; 5,880,104; 5,958,891;
15 5,985,847; 6,017,897; 6,110,898; 6,204,250, the disclosures of which are incorporated herein by reference in their entireties. For cancer therapy, it is preferred that the vector also be tumor-selective. See, e.g., Doronin *et al.*, *J. Virol.* 75: 3314-24 (2001).

In another embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a nucleic
20 acid molecule of the present invention is administered. The nucleic acid molecule can be delivered in a vector that drives expression of a CSP, fusion protein, or fragment thereof, or without such vector. Nucleic acid compositions that can drive expression of a CSP are administered, for example, to complement a deficiency in the native CSP, or as DNA vaccines. Expression vectors derived from virus, replication deficient retroviruses,
25 adenovirus, adeno-associated (AAV) virus, herpes virus, or vaccinia virus can be used as can plasmids. See, e.g., Cid-Arregui, *supra*. In a preferred embodiment, the nucleic acid molecule encodes a CSP having the amino acid sequence of SEQ ID NO: 113-259, or a fragment, fusion protein, allelic variant or homolog thereof.

In still other therapeutic methods of the present invention, pharmaceutical
30 compositions comprising host cells that express a CSP, fusions, or fragments thereof can be administered. In such cases, the cells are typically autologous, so as to circumvent xenogeneic or allotypic rejection, and are administered to complement defects in CSP

production or activity. In a preferred embodiment, the nucleic acid molecules in the cells encode a CSP having the amino acid sequence of SEQ ID NO: 113-259, or a fragment, fusion protein, allelic variant or homolog thereof.

Antisense Administration

5 Antisense nucleic acid compositions, or vectors that drive expression of a CSG antisense nucleic acid, are administered to downregulate transcription and/or translation of a CSG in circumstances in which excessive production, or production of aberrant protein, is the pathophysiologic basis of disease.

10 Antisense compositions useful in therapy can have a sequence that is complementary to coding or to noncoding regions of a CSG. For example, oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred.

Catalytic antisense compositions, such as ribozymes, that are capable of sequence-specific hybridization to CSG transcripts, are also useful in therapy. *See, e.g.*,
15 Phylactou, *Adv. Drug Deliv. Rev.* 44(2-3): 97-108 (2000); Phylactou *et al.*, *Hum. Mol. Genet.* 7(10): 1649-53 (1998); Rossi, *Ciba Found. Symp.* 209: 195-204 (1997); and Sigurdsson *et al.*, *Trends Biotechnol.* 13(8): 286-9 (1995).

Other nucleic acids useful in the therapeutic methods of the present invention are those that are capable of triplex helix formation in or near the CSG genomic locus. Such
20 triplexing oligonucleotides are able to inhibit transcription. *See, e.g.*, Intody *et al.*, *Nucleic Acids Res.* 28(21): 4283-90 (2000); and McGuffie *et al.*, *Cancer Res.* 60(14): 3790-9 (2000). Pharmaceutical compositions comprising such triplex forming oligos (TFOs) are administered in circumstances in which excessive production, or production of aberrant protein, is a pathophysiologic basis of disease.

25 In a preferred embodiment, the antisense molecule is derived from a nucleic acid molecule encoding a CSP, preferably a CSP comprising an amino acid sequence of SEQ ID NO: 113-259, or a fragment, allelic variant or homolog thereof. In a more preferred embodiment, the antisense molecule is derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-112, or a part, allelic variant, substantially similar
30 or hybridizing nucleic acid thereof.

Polypeptide Administration

In one embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a CSP, a fusion protein, fragment, analog or derivative thereof is administered to a subject with a clinically-significant CSP defect.

5 Protein compositions are administered, for example, to complement a deficiency in native CSP. In other embodiments, protein compositions are administered as a vaccine to elicit a humoral and/or cellular immune response to CSP. The immune response can be used to modulate activity of CSP or, depending on the immunogen, to immunize against aberrant or aberrantly expressed forms, such as mutant or inappropriately expressed
10 isoforms. In yet other embodiments, protein fusions having a toxic moiety are administered to ablate cells that aberrantly accumulate CSP.

In a preferred embodiment, the polypeptide administered is a CSP comprising an amino acid sequence of SEQ ID NO: 113-259, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the polypeptide
15 is encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-112, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

Antibody, Agonist and Antagonist Administration

In another embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising an antibody
20 (including fragment or derivative thereof) of the present invention is administered. As is well known, antibody compositions are administered, for example, to antagonize activity of CSP, or to target therapeutic agents to sites of CSP presence and/or accumulation. In a preferred embodiment, the antibody specifically binds to a CSP comprising an amino acid sequence of SEQ ID NO: 113-259, or a fusion protein, allelic variant, homolog, analog or
25 derivative thereof. In a more preferred embodiment, the antibody specifically binds to a CSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-112, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

The present invention also provides methods for identifying modulators which bind to a CSP or have a modulatory effect on the expression or activity of a CSP.
30 Modulators which decrease the expression or activity of CSP (antagonists) are believed to be useful in treating colon cancer. Such screening assays are known to those of skill in the art and include, without limitation, cell-based assays and cell-free assays. Small molecules

predicted via computer imaging to specifically bind to regions of a CSP can also be designed, synthesized and tested for use in the imaging and treatment of colon cancer. Further, libraries of molecules can be screened for potential anticancer agents by assessing the ability of the molecule to bind to the CSPs identified herein. Molecules identified in
5 the library as being capable of binding to a CSP are key candidates for further evaluation for use in the treatment of colon cancer. In a preferred embodiment, these molecules will downregulate expression and/or activity of a CSP in cells.

In another embodiment of the therapeutic methods of the present invention, a pharmaceutical composition comprising a non-antibody antagonist of CSP is administered.
10 Antagonists of CSP can be produced using methods generally known in the art. In particular, purified CSP can be used to screen libraries of pharmaceutical agents, often combinatorial libraries of small molecules, to identify those that specifically bind and antagonize at least one activity of a CSP.

In other embodiments a pharmaceutical composition comprising an agonist of a
15 CSP is administered. Agonists can be identified using methods analogous to those used to identify antagonists.

In a preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a CSP comprising an amino acid sequence of SEQ ID NO: 113-259, or a fusion protein, allelic variant, homolog, analog or derivative thereof.
20 In a more preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a CSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-112, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

25 *Targeting Colon Tissue*

The invention also provides a method in which a polypeptide of the invention, or an antibody thereto, is linked to a therapeutic agent such that it can be delivered to the colon or to specific cells in the colon. In a preferred embodiment, an anti-CSP antibody is linked to a therapeutic agent and is administered to a patient in need of such therapeutic
30 agent. The therapeutic agent may be a toxin, if colon tissue needs to be selectively destroyed. This would be useful for targeting and killing colon cancer cells. In another embodiment, the therapeutic agent may be a growth or differentiation factor, which would be useful for promoting colon cell function.

In another embodiment, an anti-CSP antibody may be linked to an imaging agent that can be detected using, *e.g.*, magnetic resonance imaging, CT or PET. This would be useful for determining and monitoring colon function, identifying colon cancer tumors, and identifying noncancerous colon diseases.

5

EXAMPLES

Example 1a: Alternative Splice Variants

We identified gene transcripts using the Gencarta™ tools (Compugen Ltd., Tel Aviv, Israel) and a variety of public and proprietary databases. These splice variants are either sequences which differ from a previously defined sequence or new uses of known sequences. In general related variants are annotated as DEX0449_XXX.nt.1, DEX0449_XXX.nt.2, DEX0449_XXX.nt.3, etc. The variant DNA sequences encode proteins which differ from a previously defined protein sequence. In relation to the nucleotide sequence naming convention, protein variants are annotated as DEX0449_XXX.aa.1, DEX0449_XXX.aa.2, etc., wherein transcript DEX0449_XXX.nt.1 encodes protein DEX0449_XXX.aa.1. A single transcript may encode a protein from an alternate Open Reading Frame (ORF) which is designated DEX0449_XXX.orf.1. Additionally, multiple transcripts may encode for a single protein. In this case, DEX0449_XXX.nt.1 and DEX0449_XXX.nt.2 will both be associated with DEX0449_XXX.aa.1.

The mapping of the nucleic acid ("NT") SEQ ID NO; DEX ID; chromosomal location (if known); open reading frame (ORF) location; amino acid ("AA") SEQ ID NO; AA DEX ID; are shown in the table below.

SEQ ID NO	DEX ID	Chromo Map	ORF Loc	SEQ ID NO	DEX ID
1	DEX0449_001.nt.1	17q25.3	1-498	113	DEX0449_001.aa.1
1	DEX0449_001.nt.1	17q25.3	2-493	114	DEX0449_001.orf.1
2	DEX0449_002.nt.1	5q13.1	253-1022	115	DEX0449_002.aa.1
2	DEX0449_002.nt.1	5q13.1	128-877	116	DEX0449_002.orf.1
3	DEX0449_003.nt.1	17q25.1	9-1226	117	DEX0449_003.aa.1
4	DEX0449_003.nt.2	17q25.1	7-1229	117	DEX0449_003.aa.1
5	DEX0449_003.nt.3	17q25.1	9-1583	118	DEX0449_003.aa.3
6	DEX0449_003.nt.4	17q25.1	2373-3357	119	DEX0449_003.aa.4
6	DEX0449_003.nt.4	17q25.1	2207-3358	120	DEX0449_003.orf.4

7	DEX0449_003.nt.5	17q25.1	7-1229	117	DEX0449_003.aa.1
8	DEX0449_003.nt.6	17q25.1	7-1229	117	DEX0449_003.aa.1
8	DEX0449_003.nt.6	17q25.1	2490-3839	121	DEX0449_003.orf.6
9	DEX0449_003.nt.7	17q25.1	9-914	122	DEX0449_003.aa.7
10	DEX0449_004.nt.1	4q32.3	237-722	123	DEX0449_004.aa.1
11	DEX0449_005.nt.1	8q24.3	181-802	124	DEX0449_005.aa.1
11	DEX0449_005.nt.1	8q24.3	3-1142	125	DEX0449_005.orf.1
12	DEX0449_005.nt.2	8q24.3	181-802	124	DEX0449_005.aa.1
12	DEX0449_005.nt.2	8q24.3	3-1142	126	DEX0449_005.orf.2
13	DEX0449_005.nt.3	8q24.3	229-1235	127	DEX0449_005.aa.3
13	DEX0449_005.nt.3	8q24.3	594-1574	128	DEX0449_005.orf.3
14	DEX0449_005.nt.4	8q24.3	180-662	129	DEX0449_005.aa.4
15	DEX0449_005.nt.5	8q24.3	181-667	129	DEX0449_005.aa.4
16	DEX0449_005.nt.6	8q24.3	181-802	124	DEX0449_005.aa.1
16	DEX0449_005.nt.6	8q24.3	3-980	130	DEX0449_005.orf.6
17	DEX0449_006.nt.1	14q32.33	1774-2421	131	DEX0449_006.aa.1
18	DEX0449_007.nt.1	4q13.3	103-431	132	DEX0449_007.aa.1
18	DEX0449_007.nt.1	4q13.3	2-427	133	DEX0449_007.orf.1
19	DEX0449_008.nt.1	4q21.1	232-1681	134	DEX0449_008.aa.1
19	DEX0449_008.nt.1	4q21.1	504-1679	135	DEX0449_008.orf.1
20	DEX0449_009.nt.1	4q21.1	1-951	136	DEX0449_009.aa.1
20	DEX0449_009.nt.1	4q21.1	3-944	137	DEX0449_009.orf.1
21	DEX0449_009.nt.2	4q21.1	48-554	138	DEX0449_009.aa.2
22	DEX0449_010.nt.1	15q25.3	117-561	139	DEX0449_010.aa.1
22	DEX0449_010.nt.1	15q25.3	61-558	140	DEX0449_010.orf.1
23	DEX0449_011.nt.1	8q11.23	868-1633	141	DEX0449_011.aa.1
23	DEX0449_011.nt.1	8q11.23	723-1628	142	DEX0449_011.orf.1
24	DEX0449_012.nt.1	7p11.2	733-1407	143	DEX0449_012.aa.1
25	DEX0449_012.nt.2	7p11.2	918-1596	143	DEX0449_012.aa.1
26	DEX0449_012.nt.3	7p11.2	733-1411	143	DEX0449_012.aa.1
27	DEX0449_013.nt.1	8p21.2	66-812	144	DEX0449_013.aa.1
28	DEX0449_014.nt.1	18p11.31	880-1218	145	DEX0449_014.aa.1
29	DEX0449_015.nt.1	20q13.2	422-727	146	DEX0449_015.aa.1
30	DEX0449_016.nt.1	7p22.3	394-1633	147	DEX0449_016.aa.1
30	DEX0449_016.nt.1	7p22.3	1019-1627	148	DEX0449_016.orf.1
31	DEX0449_017.nt.1	17p13.3	295-642	149	DEX0449_017.aa.1
31	DEX0449_017.nt.1	17p13.3	138-560	150	DEX0449_017.orf.1
32	DEX0449_018.nt.1	2q35	133-1414	151	DEX0449_018.aa.1

33	DEX0449_018.nt.2	2q35	133-1414	151	DEX0449_018.aa.1
34	DEX0449_018.nt.3	2q35	133-1414	151	DEX0449_018.aa.1
35	DEX0449_018.nt.4	2q35	134-1243	152	DEX0449_018.aa.4
36	DEX0449_018.nt.5	2q35	134-757	153	DEX0449_018.aa.5
37	DEX0449_018.nt.6	2q35	133-1414	151	DEX0449_018.aa.1
38	DEX0449_018.nt.7	2q35	2635-3261	154	DEX0449_018.orf.7
38	DEX0449_018.nt.7	2q35	133-511	155	DEX0449_018.aa.7
39	DEX0449_018.nt.8	2q35	134-1411	151	DEX0449_018.aa.1
40	DEX0449_019.nt.1	9	2070-2642	156	DEX0449_019.orf.1
40	DEX0449_019.nt.1	9	1311-1704	157	DEX0449_019.aa.1
41	DEX0449_020.nt.1	19q13.41	266-1285	158	DEX0449_020.orf.1
41	DEX0449_020.nt.1	19q13.41	364-1286	159	DEX0449_020.aa.1
42	DEX0449_020.nt.2	19q13.41	3-1457	160	DEX0449_020.orf.2
42	DEX0449_020.nt.2	19q13.41	280-1679	161	DEX0449_020.aa.2
43	DEX0449_021.nt.1	1q32.1	183-404	162	DEX0449_021.orf.1
43	DEX0449_021.nt.1	1q32.1	1-193	163	DEX0449_021.aa.1
44	DEX0449_022.nt.1	17q25.3	612-1550	164	DEX0449_022.aa.1
45	DEX0449_022.nt.2	17q25.3	72-1256	165	DEX0449_022.orf.2
45	DEX0449_022.nt.2	17q25.3	2750-3608	166	DEX0449_022.aa.2
46	DEX0449_023.nt.1	16p13.2	1330-1839	167	DEX0449_023.orf.1
46	DEX0449_023.nt.1	16p13.2	1362-1840	168	DEX0449_023.aa.1
47	DEX0449_023.nt.2	16p13.2	1108-1617	169	DEX0449_023.orf.2
47	DEX0449_023.nt.2	16p13.2	1140-1618	168	DEX0449_023.aa.1
48	DEX0449_024.nt.1	16p13.2	502-1266	170	DEX0449_024.orf.1
48	DEX0449_024.nt.1	16p13.2	590-1034	171	DEX0449_024.aa.1
49	DEX0449_024.nt.2	16p13.2	264-1028	172	DEX0449_024.orf.2
49	DEX0449_024.nt.2	16p13.2	352-796	171	DEX0449_024.aa.1
50	DEX0449_024.nt.3	16p13.2	2-730	173	DEX0449_024.orf.3
50	DEX0449_024.nt.3	16p13.2	55-498	174	DEX0449_024.aa.3
51	DEX0449_024.nt.4	16p13.2	948-1553	175	DEX0449_024.orf.4
51	DEX0449_024.nt.4	16p13.2	1104-1521	176	DEX0449_024.aa.4
52	DEX0449_024.nt.5	16p13.2	232-789	177	DEX0449_024.aa.5
52	DEX0449_024.nt.5	16p13.2	30-824	178	DEX0449_024.orf.5
53	DEX0449_024.nt.6	16p13.2	606-1208	179	DEX0449_024.aa.6
54	DEX0449_025.nt.1	3q25.32	93-569	180	DEX0449_025.aa.1
55	DEX0449_025.nt.2	3q25.32	454-837	181	DEX0449_025.aa.2
56	DEX0449_026.nt.1	5q35.1	90-761	182	DEX0449_026.aa.1

57	DEX0449_026.nt.2	5q35.1	1108-1801	183	DEX0449_026.aa.2
57	DEX0449_026.nt.2	5q35.1	1123-1797	184	DEX0449_026.orf.2
58	DEX0449_027.nt.1	19p13.3	157-3416	185	DEX0449_027.aa.1
58	DEX0449_027.nt.1	19p13.3	3-2006	186	DEX0449_027.orf.1
59	DEX0449_027.nt.2	19p13.3	231-2543	187	DEX0449_027.aa.2
60	DEX0449_027.nt.3	19p13.3	159-2702	188	DEX0449_027.aa.3
61	DEX0449_028.nt.1	8q13.2	152-523	189	DEX0449_028.aa.1
62	DEX0449_029.nt.1	1q21.3	187-1074	190	DEX0449_029.aa.1
63	DEX0449_030.nt.1	20q11.23	1-589	191	DEX0449_030.aa.1
63	DEX0449_030.nt.1	20q11.23	3-584	192	DEX0449_030.orf.1
64	DEX0449_031.nt.1	20q13.33	1827-2226	193	DEX0449_031.aa.1
64	DEX0449_031.nt.1	20q13.33	1602-2198	194	DEX0449_031.orf.1
65	DEX0449_032.nt.1	11q11	133-913	195	DEX0449_032.aa.1
66	DEX0449_032.nt.2	11q11	252-723	196	DEX0449_032.aa.2
66	DEX0449_032.nt.2	11q11	212-715	197	DEX0449_032.orf.2
67	DEX0449_032.nt.3	11q11	134-544	198	DEX0449_032.aa.3
68	DEX0449_032.nt.4	11q11	97-810	199	DEX0449_032.aa.4
68	DEX0449_032.nt.4	11q11	341-808	200	DEX0449_032.orf.4
69	DEX0449_033.nt.1	22q13.31	303-566	201	DEX0449_033.aa.1
70	DEX0449_033.nt.2	22q13.31	43-166	202	DEX0449_033.aa.2
70	DEX0449_033.nt.2	22q13.31	33-212	203	DEX0449_033.orf.2
71	DEX0449_034.nt.1	19p13.2	34-672	204	DEX0449_034.aa.1
72	DEX0449_035.nt.1	3q13.11	1-1378	205	DEX0449_035.aa.1
72	DEX0449_035.nt.1	3q13.11	60-1376	206	DEX0449_035.orf.1
73	DEX0449_036.nt.1	1q42.12	79-473	207	DEX0449_036.aa.1
73	DEX0449_036.nt.1	1q42.12	17-745	208	DEX0449_036.orf.1
74	DEX0449_037.nt.1	7p15.3	625-1236	209	DEX0449_037.aa.1
75	DEX0449_038.nt.1	20q13.13	2-241	210	DEX0449_038.aa.1
76	DEX0449_038.nt.2	20q13.13	1-244	210	DEX0449_038.aa.1
76	DEX0449_038.nt.2	20q13.13	669-920	211	DEX0449_038.orf.2
77	DEX0449_038.nt.3	20q13.13	1-168	212	DEX0449_038.aa.3
78	DEX0449_038.nt.4	20q13.13	1-244	210	DEX0449_038.aa.1
79	DEX0449_039.nt.1	1p34.1	987-2339	213	DEX0449_039.aa.1
80	DEX0449_040.nt.1	6p25.3	272-1282	214	DEX0449_040.aa.1
81	DEX0449_040.nt.2	6p25.3	272-1267	215	DEX0449_040.aa.2
82	DEX0449_040.nt.3	6p25.3	272-1417	216	DEX0449_040.aa.3
83	DEX0449_041.nt.1	19	1892-2665	217	DEX0449_041.aa.1
84	DEX0449_041.nt.2	19	1653-2442	218	DEX0449_041.aa.2

84	DEX0449_041.nt.2	19	1620-2492	219	DEX0449_041.orf.2
85	DEX0449_041.nt.3	19	1891-2593	220	DEX0449_041.aa.3
85	DEX0449_041.nt.3	19	1475-2644	221	DEX0449_041.orf.3
86	DEX0449_042.nt.1	17q11.1	3-332	222	DEX0449_042.aa.1
87	DEX0449_043.nt.1	X;22884428-22906568	263-1033	223	DEX0449_043.aa.1
88	DEX0449_044.nt.1	5p13.1	143-461	224	DEX0449_044.aa.1
88	DEX0449_044.nt.1	5p13.1	125-457	225	DEX0449_044.orf.1
89	DEX0449_044.nt.2	5p13.1	191-506	226	DEX0449_044.aa.2
89	DEX0449_044.nt.2	5p13.1	16-402	227	DEX0449_044.orf.2
90	DEX0449_044.nt.3	5p13.1	3-290	228	DEX0449_044.aa.3
91	DEX0449_044.nt.4	5p13.1	1-168	229	DEX0449_044.aa.4
91	DEX0449_044.nt.4	5p13.1	51-266	230	DEX0449_044.orf.4
92	DEX0449_045.nt.1	19q13.41	5-563	231	DEX0449_045.aa.1
92	DEX0449_045.nt.1	19q13.41	462-1103	232	DEX0449_045.orf.1
93	DEX0449_045.nt.2	16q23.1	5-398	233	DEX0449_045.aa.2
93	DEX0449_045.nt.2	16q23.1	1-396	234	DEX0449_045.orf.2
94	DEX0449_045.nt.3	19q13.41	10-594	235	DEX0449_045.aa.3
95	DEX0449_046.nt.1	X;46271396-46272517	3-347	236	DEX0449_046.aa.1
96	DEX0449_046.nt.2	X;46271094-46272517	2-469	237	DEX0449_046.aa.2
97	DEX0449_047.nt.1	Un_1;3772342-3781437	19-696	238	DEX0449_047.aa.1
97	DEX0449_047.nt.1	Un_1;3772342-3781437	2470-3228	239	DEX0449_047.orf.1
98	DEX0449_047.nt.2	Un_1;3772342-3781437	19-699	238	DEX0449_047.aa.1
98	DEX0449_047.nt.2	Un_1;3772342-3781437	2470-3507	240	DEX0449_047.orf.2
99	DEX0449_048.nt.1	11p15.5	1-277	241	DEX0449_048.aa.1
99	DEX0449_048.nt.1	11p15.5	91-366	242	DEX0449_048.orf.1
100	DEX0449_049.nt.1	3q21.3	34-447	243	DEX0449_049.aa.1
100	DEX0449_049.nt.1	3q21.3	1-444	244	DEX0449_049.orf.1
101	DEX0449_050.nt.1	3q21.3	86-1522	245	DEX0449_050.aa.1
102	DEX0449_050.nt.2	3q21.3	86-1168	246	DEX0449_050.aa.2
103	DEX0449_050.nt.3	3q21.3	86-1345	247	DEX0449_050.aa.3
104	DEX0449_051.nt.1	5q35.3	2-385	248	DEX0449_051.aa.1
105	DEX0449_052.nt.1	13q34	141-1085	249	DEX0449_052.aa.1
106	DEX0449_053.nt.1	7q36.3	674-1099	250	DEX0449_053.aa.1
107	DEX0449_053.nt.2	7q36.3	402-621	251	DEX0449_053.aa.2
107	DEX0449_053.nt.2	7q36.3	622-987	252	DEX0449_053.orf.2
108	DEX0449_054.nt.1	9q34.11	64-194	253	DEX0449_054.aa.1

108	DEX0449_054.nt.1	9q34.11	895-1674	254	DEX0449_054.orf.1
109	DEX0449_054.nt.2	9q34.11	64-194	253	DEX0449_054.aa.1
109	DEX0449_054.nt.2	9q34.11	195-1049	255	DEX0449_054.orf.2
110	DEX0449_055.nt.1	19p13.3	87-1890	256	DEX0449_055.aa.1
110	DEX0449_055.nt.1	19p13.3	28-1887	257	DEX0449_055.orf.1
111	DEX0449_055.nt.2	19p13.3	26-736	258	DEX0449_055.aa.2
112	DEX0449_055.nt.3	19p13.3	87-1890	256	DEX0449_055.aa.1
112	DEX0449_055.nt.3	19p13.3	28-1887	259	DEX0449_055.orf.3

The polypeptides of the present invention were analyzed and the following attributes were identified; specifically, epitopes, post translational modifications, signal peptides and transmembrane domains. Antigenicity (Epitope) prediction was performed through the antigenic module in the EMBOSS package. Rice, P., EMBOSS: The European Molecular Biology Open Software Suite, *Trends in Genetics* 16(6): 276-277 (2000). The antigenic module predicts potentially antigenic regions of a protein sequence, using the method of Kolaskar and Tongaonkar. Kolaskar, AS and Tongaonkar, PC., A semi-empirical method for prediction of antigenic determinants on protein antigens, *FEBS Letters* 276: 172-174 (1990). Examples of post-translational modifications (PTMs) and other motifs of the CSPs of this invention are listed below. In addition, antibodies that specifically bind such post-translational modifications may be useful as a diagnostic or as therapeutic. The PTMs and other motifs were predicted by using the ProSite Dictionary of Proteins Sites and Patterns (Bairoch *et al.*, *Nucleic Acids Res.* 25(1):217-221 (1997)), the following motifs, including PTMs, were predicted for the CSPs of the invention. The signal peptides were detected by using the SignalP 2.0, *see Nielsen et al.*, *Protein Engineering* 12, 3-9 (1999). Prediction of transmembrane helices in proteins was performed by the application TMHMM 2.0, "currently the best performing transmembrane prediction program", according to authors (Krogh *et al.*, *Journal of Molecular Biology*, 305(3):567-580, (2001); Moller *et al.*, *Bioinformatics*, 17(7):646-653, (2001); Sonnhammer, *et al.*, *A hidden Markov model for predicting transmembrane helices in protein sequences* in Glasgow, *et al.* Ed. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, pages 175-182, Menlo Park, CA, 1998. AAAI Press. The PSORT II program may also be used to predict cellular localizations. Horton *et al.*, *Intelligent Systems for Molecular Biology* 5: 147-152 (1997).

The table below includes the following sequence annotations: Signal peptide presence; TM (number of membrane domain, topology in orientation and position); Amino acid location and antigenic index (location, AI score); PTM and other motifs (type, amino acid residue locations); and functional domains (type, amino acid residue locations).

DEX ID	Sig P	TMHMM	Antigenicity	PTM	Domains
DEX0449_00 1.aa.1	N	0 - 01- 165;	51-59,1.044; 98- 105,1.041; 124- 130,1.064; 65-70,1.029; 143- 162,1.163; 83-96,1.122; 108- 116,1.104;	MYRISTYL 9-14; MYRISTYL 44-49; MYRISTYL 12-17; CK2_PHOSPHO_SITE 94-97; MYRISTYL 73-78; MYRISTYL 40-45; MYRISTYL 77-82; PKC_PHOSPHO_SITE 118-120; MYRISTYL 17-22; PKC_PHOSPHO_SITE 107-109; CK2_PHOSPHO_SITE 87-90; MYRISTYL 92-97; MYRISTYL 16-21;	RRM 82-152; rrm 84- 154; RRM 83-155; GLY_RICH 1-82;
DEX0449_00 1.orf.1	N	0 - 01- 164;		MYRISTYL 3-8; MYRISTYL 72-77; MYRISTYL 15-20; MYRISTYL 91-96; PKC_PHOSPHO_SITE 106-108; CK2_PHOSPHO_SITE 86-89; MYRISTYL 39-44; MYRISTYL 16-21; MYRISTYL 43-48; MYRISTYL 11-16; CK2_PHOSPHO_SITE 93-96; MYRISTYL 76-81; PKC_PHOSPHO_SITE 117-119;	RRM 82-154; RRM 81- 151; GLY_RICH 1-81; rrm 83-153;
DEX0449_00 2.aa.1	N	0 - 01- 256;	134- 156,1.159; 116- 125,1.111; 24-34,1.087; 186- 196,1.13; 71-101,1.22; 203- 226,1.227; 164- 180,1.143; 237- 245,1.093;	PKC_PHOSPHO_SITE 71-73; MYRISTYL 199-204; CK2_PHOSPHO_SITE 128-131; MYRISTYL 49-54; MYRISTYL 233-238; MYRISTYL 134-139; PKC_PHOSPHO_SITE 34-36; MYRISTYL 161-166;	PRO_RICH 80-129;
DEX0449_00	Y	0 - 01-	158-	MYRISTYL 176-181;	PRO_RICH 122-171;

2.orf.1		250;	167,1.111; 66-76,1.087; 113- 143,1.22; 176- 198,1.159; 9-36,1.178; 206- 222,1.143; 228- 238,1.13;	PKC_PHOSPHO_SITE 76-78; MYRISTYL 91-96; PKC_PHOSPHO_SITE 4-6; MYRISTYL 241-246; MYRISTYL 203-208; CK2_PHOSPHO_SITE 170-173; PKC_PHOSPHO_SITE 3-5; PKC_PHOSPHO_SITE 113-115;	
DEX0449_00 3.aa.1	N	0 - 01- 406;	82-97,1.113; 201- 207,1.14; 59-75,1.139; 124- 130,1.123; 213- 226,1.126; 142- 174,1.207; 356- 375,1.115; 251- 278,1.154; 388- 403,1.127; 334- 340,1.063; 323- 332,1.205; 182- 191,1.186; 105- 121,1.136; 4-11,1.111; 239- 248,1.148; 289- 300,1.104; 20-26,1.083;	MYRISTYL 335-340; CK2_PHOSPHO_SITE 30-33; CK2_PHOSPHO_SITE 49-52; PKC_PHOSPHO_SITE 120-122; CK2_PHOSPHO_SITE 89-92; MYRISTYL 198-203; PKC_PHOSPHO_SITE 13-15; PKC_PHOSPHO_SITE 49-51; CK2_PHOSPHO_SITE 88-91; PKC_PHOSPHO_SITE 79-81; PKC_PHOSPHO_SITE 344-346; MYRISTYL 189-194; ASN_GLYCOSYLATION 379-382;	Glyco_hydro_85 131- 385;
DEX0449_00 3.aa.3	N	0 - 01- 525;	450- 457,1.192; 411- 418,1.091; 124- 130,1.123; 347- 373,1.136; 20-26,1.083; 490- 502,1.207; 182- 191,1.186; 289- 300,1.104; 213-	PKC_PHOSPHO_SITE 79-81; MYRISTYL 198-203; CK2_PHOSPHO_SITE 88-91; MYRISTYL 189-194; PKC_PHOSPHO_SITE 344-346; PKC_PHOSPHO_SITE 120-122; PKC_PHOSPHO_SITE 13-15; MYRISTYL 335-340; MYRISTYL 486-491; CK2_PHOSPHO_SITE 49-52;	BRCT 337-383; Glyco_hydro_85 131- 414;

			226,1.126; 105- 121,1.136; 239- 248,1.148; 251- 278,1.154; 142- 174,1.207; 82-97,1.113; 59-75,1.139; 334- 340,1.063; 462- 483,1.166; 422- 438,1.104; 323- 332,1.205; 201- 207,1.14; 4- 11,1.111; 387- 408,1.195;	CK2_PHOSPHO_SITE 428-431; PKC_PHOSPHO_SITE 49-51; CK2_PHOSPHO_SITE 30-33; MYRISTYL 416-421; CK2_PHOSPHO_SITE 89-92;	
DEX0449_00 3.aa.4	N	0 - 01- 327;	11-18,1.089; 270- 289,1.165; 90-97,1.156; 134- 143,1.053; 29-36,1.1; 60-86,1.143; 44-57,1.084; 317- 324,1.169; 223- 250,1.179; 120- 132,1.108; 146- 171,1.237; 298- 314,1.155; 182- 214,1.168; 102-115,1.1;	CK2_PHOSPHO_SITE 175-178; MYRISTYL 37-42; CK2_PHOSPHO_SITE 216-219; MYRISTYL 272-277; PKC_PHOSPHO_SITE 182-184; PKC_PHOSPHO_SITE 218-220; CK2_PHOSPHO_SITE 218-221; PKC_PHOSPHO_SITE 122-124; PKC_PHOSPHO_SITE 156-158; CK2_PHOSPHO_SITE 98-101; PKC_PHOSPHO_SITE 118-120; MYRISTYL 9-14;	
DEX0449_00 3.orf.4	N	0 - 01- 384;	119- 143,1.143; 355- 371,1.155; 280- 307,1.179; 177- 189,1.108; 95- 112,1.089; 327- 346,1.165; 374-	PKC_PHOSPHO_SITE 90-92; PKC_PHOSPHO_SITE 8-10; CK2_PHOSPHO_SITE 273-276; PKC_PHOSPHO_SITE 175-177; CK2_PHOSPHO_SITE 232-235; PKC_PHOSPHO_SITE 179-181; MYRISTYL 329-334;	

			381,1.169; 85-90,1.061; 239- 271,1.168; 191- 200,1.053; 159-172,1.1; 203- 228,1.237; 147- 154,1.156; 4-64,1.135;	CK2_PHOSPHO_SITE 155-158; CK2_PHOSPHO_SITE 275-278; MYRISTYL 94-99; PKC_PHOSPHO_SITE 239-241; MYRISTYL 81-86; PKC_PHOSPHO_SITE 275-277; CK2_PHOSPHO_SITE 90-93; MYRISTYL 47-52; PKC_PHOSPHO_SITE 213-215;	
DEX0449_00 3.orf.6	N	0 - 01- 450;	147- 154,1.156; 95- 112,1.089; 255- 269,1.178; 159-172,1.1; 421- 437,1.155; 292- 337,1.244; 4-64,1.135; 440- 447,1.169; 239- 250,1.08; 393- 412,1.165; 177- 189,1.108; 191- 200,1.053; 85-90,1.061; 119- 143,1.143; 346- 373,1.179; 271- 278,1.063; 203- 228,1.237;	CK2_PHOSPHO_SITE 155-158; PKC_PHOSPHO_SITE 90-92; CK2_PHOSPHO_SITE 90-93; CK2_PHOSPHO_SITE 232-235; PKC_PHOSPHO_SITE 8-10; MYRISTYL 47-52; MYRISTYL 252-257; MYRISTYL 287-292; PKC_PHOSPHO_SITE 179-181; MYRISTYL 395-400; PKC_PHOSPHO_SITE 341-343; CK2_PHOSPHO_SITE 341-344; PKC_PHOSPHO_SITE 213-215; CK2_PHOSPHO_SITE 339-342; PKC_PHOSPHO_SITE 239-241; MYRISTYL 81-86; MYRISTYL 94-99; PKC_PHOSPHO_SITE 175-177;	
DEX0449_00 3.aa.7	N	0 - 01- 302;	82-97,1.113; 292- 299,1.169; 273- 289,1.155; 20-26,1.083; 124- 130,1.123; 105- 121,1.136; 239- 248,1.148; 182-	PKC_PHOSPHO_SITE 79-81; CK2_PHOSPHO_SITE 30-33; MYRISTYL 189-194; PKC_PHOSPHO_SITE 120-122; MYRISTYL 198-203; CK2_PHOSPHO_SITE 49-52; PKC_PHOSPHO_SITE 49-51; CK2 PHOSPHO SITE	Glyco_hydro_85 131- 297;

			191,1.186; 251- 264,1.138; 201- 207,1.14; 213- 226,1.126; 59-75,1.139; 4-11,1.111; 142- 174,1.207;	88-91; PKC_PHOSPHO_SITE 13-15; CK2_PHOSPHO_SITE 89-92;	
DEX0449_00 4.aa.1	Y	1 - o1- 62;tm63 - 85;i86- 162;	4-19,1.198; 55-67,1.104; 22-48,1.14; 69- 116,1.236;	PKC_PHOSPHO_SITE 87-89; CK2_PHOSPHO_SITE 98-101; CK2_PHOSPHO_SITE 135-138; MYRISTYL 128-133;	Sterol_desat 1-129; SUR2_DOMAIN 18-126;
DEX0449_00 5.aa.1	N	0 - o1- 206;	186- 197,1.16; 36-45,1.101; 97- 119,1.199; 163- 173,1.076; 142- 147,1.042; 8-31,1.142; 58-76,1.16;	CK2_PHOSPHO_SITE 100-103; PKC_PHOSPHO_SITE 7-9; CK2_PHOSPHO_SITE 42-45; PKC_PHOSPHO_SITE 194-196; PKC_PHOSPHO_SITE 42-44; MYRISTYL 75-80; PKC_PHOSPHO_SITE 183-185;	
DEX0449_00 5.orf.1	N	0 - o1- 380;	156- 178,1.199; 349- 355,1.062; 243- 259,1.168; 117- 135,1.16; 95- 104,1.101; 4-10,1.175; 275- 293,1.148; 360- 371,1.088; 43-54,1.098; 315- 322,1.094; 217- 235,1.131; 296- 304,1.127; 17-31,1.101; 261- 267,1.072; 201- 206,1.042; 67-90,1.142;	MYRISTYL 261-266; PKC_PHOSPHO_SITE 101-103; CK2_PHOSPHO_SITE 40-43; MYRISTYL 52-57; PKC_PHOSPHO_SITE 310-312; PKC_PHOSPHO_SITE 332-334; CK2_PHOSPHO_SITE 333-336; CK2_PHOSPHO_SITE 101-104; PKC_PHOSPHO_SITE 66-68; CK2_PHOSPHO_SITE 310-313; CK2_PHOSPHO_SITE 302-305; MYRISTYL 134-139; CK2_PHOSPHO_SITE 159-162; PKC_PHOSPHO_SITE 9-11; MYRISTYL 221-226;	

DEX0449_00 5.orf.2	N	0 - 01- 380;	217- 235,1.131; 296- 304,1.127; 156- 178,1.199; 95- 104,1.101; 349- 355,1.062; 67-90,1.142; 17-31,1.101; 201- 206,1.042; 243- 259,1.168; 261- 267,1.072; 43-54,1.098; 315- 322,1.094; 4-10,1.175; 275- 293,1.148; 360- 371,1.088; 117- 135,1.16;	PKC_PHOSPHO_SITE 101-103; PKC_PHOSPHO_SITE 66-68; CK2_PHOSPHO_SITE 310-313; CK2_PHOSPHO_SITE 302-305; MYRISTYL 261-266; PKC_PHOSPHO_SITE 310-312; MYRISTYL 134-139; CK2_PHOSPHO_SITE 159-162; MYRISTYL 52-57; PKC_PHOSPHO_SITE 9-11; CK2_PHOSPHO_SITE 101-104; CK2_PHOSPHO_SITE 40-43; CK2_PHOSPHO_SITE 333-336; PKC_PHOSPHO_SITE 332-334; MYRISTYL 221-226;	
DEX0449_00 5.aa.3	N	0 - 01- 334;	94- 105,1.164; 46-52,1.071; 114- 124,1.143; 270- 275,1.042; 225- 247,1.199; 314- 325,1.16; 186- 204,1.16; 291- 301,1.076; 136- 159,1.142; 10-16,1.095; 18-37,1.212; 164- 173,1.101; 76-86,1.121; 59-65,1.085;	CAMP_PHOSPHO_SITE 73-76; MYRISTYL 95-100; PKC_PHOSPHO_SITE 135-137; CK2_PHOSPHO_SITE 170-173; MYRISTYL 85-90; PKC_PHOSPHO_SITE 170-172; PKC_PHOSPHO_SITE 322-324; MYRISTYL 71-76; MYRISTYL 51-56; MYRISTYL 31-36; PKC_PHOSPHO_SITE 311-313; PKC_PHOSPHO_SITE 28-30; MYRISTYL 82-87; MYRISTYL 46-51; MYRISTYL 203-208; CK2_PHOSPHO_SITE 228-231;	
DEX0449_00 5.orf.3	N	0 - 01- 327;	307- 318,1.088; 64-82,1.16; 243- 251,1.127; 262- 269,1.094;	MYRISTYL 168-173; CK2_PHOSPHO_SITE 257-260; MYRISTYL 81-86; PKC_PHOSPHO_SITE 257-259; CK2 PHOSPHO SITE	

			164- 182,1.131; 148- 153,1.042; 42-51,1.101; 296- 302,1.062; 222- 240,1.148; 14-37,1.142; 208- 214,1.072; 190- 206,1.168; 103- 125,1.199;	280-283; MYRISTYL 208-213; CK2_PHOSPHO_SITE 48-51; CK2_PHOSPHO_SITE 106-109; CK2_PHOSPHO_SITE 3-6; CK2_PHOSPHO_SITE 249-252; PKC_PHOSPHO_SITE 48-50; PKC_PHOSPHO_SITE 279-281; PKC_PHOSPHO_SITE 13-15;	
DEX0449_00 5.aa.4	N	0 - ol- 161;	36-45,1.101; 142- 147,1.042; 58-76,1.16; 97- 119,1.199; 8-31,1.142;	PKC_PHOSPHO_SITE 42-44; ASN_GLYCOSYLATION 154-157; MYRISTYL 75-80; CK2_PHOSPHO_SITE 42-45; PKC_PHOSPHO_SITE 156-158; PKC_PHOSPHO_SITE 7-9; CK2_PHOSPHO_SITE 100-103;	
DEX0449_00 5.orf.6	N	0 - ol- 326;	156- 178,1.199; 217- 235,1.131; 43-54,1.098; 67-90,1.142; 296- 314,1.123; 261- 267,1.072; 117- 135,1.16; 275- 293,1.148; 201- 206,1.042; 243- 259,1.168; 17-31,1.101; 4-10,1.175; 95- 104,1.101;	MYRISTYL 261-266; PKC_PHOSPHO_SITE 9-11; MYRISTYL 221-226; PKC_PHOSPHO_SITE 66-68; CK2_PHOSPHO_SITE 101-104; MYRISTYL 134-139; MYRISTYL 52-57; PKC_PHOSPHO_SITE 101-103; CK2_PHOSPHO_SITE 40-43; CK2_PHOSPHO_SITE 159-162;	
DEX0449_00 6.aa.1	N	0 - ol- 216;	23-80,1.213; 88-99,1.097; 123- 140,1.268; 9-15,1.084; 109- 116,1.12; 142-	CK2_PHOSPHO_SITE 89-92; PKC_PHOSPHO_SITE 99-101; MYRISTYL 84-89; CAMP_PHOSPHO_SITE 79-82;	PRO_RICH 31-209;

			207,1.19;		
DEX0449_00 7.aa.1	Y	0 - ol- 108;	4-10,1.067; 46-52,1.055; 12-30,1.269; 59-105,1.18; 39-44,1.054;	PKC_PHOSPHO_SITE 36-38; PKC_PHOSPHO_SITE 29-31;	SMALLCYTKCXC 75-83; SMALLCYTKCXC 90- 108;
DEX0449_00 7.orf.1	Y	0 - ol- 142;	67-73,1.058; 46-58,1.043; 93-139,1.18; 4-30,1.163; 37-44,1.048;	PKC_PHOSPHO_SITE 10-12; MYRISTYL 7-12;	SMALLCYTKCXC 124- 142; INTERLEUKIN8 125-142; INTERLEUKIN8 102- 125; SMALLCYTKCXC 109-117;
DEX0449_00 8.aa.1	N	0 - ol- 482;	298- 306,1.134; 176- 183,1.113; 16-33,1.107; 325- 341,1.104; 38-45,1.113; 373- 394,1.127; 285- 294,1.256; 257- 275,1.079; 354- 361,1.095; 230- 254,1.098; 52-58,1.067; 121- 128,1.113; 7-14,1.136; 110- 119,1.111; 62- 101,1.184;	MYRISTYL 44-49; ASN_GLYCOSYLATION 317-320; PKC_PHOSPHO_SITE 48-50; ASN_GLYCOSYLATION 285-288; MYRISTYL 43-48; CK2_PHOSPHO_SITE 141-144; TYR_PHOSPHO_SITE 50-56; MYRISTYL 96-101; ASN_GLYCOSYLATION 69-72; MYRISTYL 258-263; MYRISTYL 471-476; ASN_GLYCOSYLATION 37-40; PKC_PHOSPHO_SITE 250-252; CK2_PHOSPHO_SITE 204-207; MYRISTYL 435-440; CK2_PHOSPHO_SITE 85-88; MYRISTYL 465-470; CK2_PHOSPHO_SITE 192-195; PKC_PHOSPHO_SITE 466-468; MYRISTYL 118-123; MYRISTYL 439-444; CK2_PHOSPHO_SITE 149-152;	RRM 332-405; RRM 331-409; NTF2 11- 133; GLY_RICH 419- 479; NTF2_DOMAIN 11-133; rrm 333- 404; sp_Q9UPA1_Q9UPA1_HU MAN 11-132; GLU_RICH 134-223;
DEX0449_00 8.orf.1	N	0 - ol- 392;	4-12,1.235; 264- 271,1.095; 283- 304,1.127; 235- 251,1.104; 86-93,1.113; 37-45,1.123; 140- 165.1.098;	CK2_PHOSPHO_SITE 59-62; RGD 221- 223; MYRISTYL 168-173; CK2_PHOSPHO_SITE 114-117; CK2_PHOSPHO_SITE 102-105; MYRISTYL 345-350; PKC_PHOSPHO_SITE 376-378;	RRM 242-315; HUDSXL RNA 246-261; HUDSXL RNA 261-273; rrm 243-314; PRICHEXTENSIN 101- 122; PRICHEXTENSIN 194-219; GLU_RICH 52-133; PRICHEXTENSIN 164- 181; GLY_RICH 329- 389; RRM 241-319;

			16-33,1.144; 167- 201,1.142;	PKC_PHOSPHO_SITE 160-162; MYRISTYL 375-380; ASN_GLYCOSYLATION 227-230; MYRISTYL 349-354; CK2_PHOSPHO_SITE 51-54; MYRISTYL 381-386;	PRICHEXTENSN 136- 152; PRO_RICH 138- 219;
DEX0449_00 9.aa.1	N	0 - o1- 316;	159- 175,1.104; 233- 239,1.056; 94- 107,1.098; 207- 228,1.127; 109- 128,1.075; 30-37,1.065; 5-15,1.225; 188- 195,1.095; 41-50,1.113; 135- 140,0.986;	CK2_PHOSPHO_SITE 71-74; CK2_PHOSPHO_SITE 16-19; MYRISTYL 269-274; ASN_GLYCOSYLATION 151-154; MYRISTYL 305-310; PKC_PHOSPHO_SITE 300-302; CK2_PHOSPHO_SITE 14-17; MYRISTYL 299-304; MYRISTYL 273-278; RGD 145- 147; CK2_PHOSPHO_SITE 59-62;	RRM 166-239; rrm 167-238; HUDSXL RNA 170-185; HUDSXL RNA 185-197; GLU_RICH 17-90; RRM 165-243; PRO_RICH 95-143; GLY_RICH 253-313;
DEX0449_00 9.orf.1	N	0 - o1- 314;		MYRISTYL 267-272; MYRISTYL 271-276; ASN_GLYCOSYLATION 149-152; MYRISTYL 2-7; ASN_GLYCOSYLATION 4-7; RGD 143-145; PKC_PHOSPHO_SITE 298-300; CK2_PHOSPHO_SITE 57-60; MYRISTYL 297-302; CK2_PHOSPHO_SITE 14-17; MYRISTYL 303-308; CK2_PHOSPHO_SITE 12-15; CK2_PHOSPHO_SITE 69-72;	HUDSXL RNA 183-195; GLU_RICH 15-88; PRO_RICH 93-141; RRM 164-237; rrm 165-236; GLY_RICH 251-311; HUDSXL RNA 168-183; RRM 163- 241;
DEX0449_00 9.aa.2	N	0 - i1- 169;	41-48,1.095; 125- 130,1.013; 11-28,1.104; 86-92,1.056; 55-82,1.127;	MYRISTYL 126-131; MYRISTYL 122-127; MYRISTYL 152-157; MYRISTYL 158-163; ASN_GLYCOSYLATION 4-7; PKC_PHOSPHO_SITE 153-155;	RRM 19-92; HUDSXL RNA 23-38; GLY_RICH 106-166; RRM 18-96; HUDSXL RNA 38-50; rrm 20-91;
DEX0449_01 0.aa.1	N	0 - i1- 147;	85-99,1.144; 9-23,1.113; 66-73,1.073; 51-59,1.107; 106-	PKC_PHOSPHO_SITE 43-45; PKC_PHOSPHO_SITE 8-10; MYRISTYL 112-117;	Ribosomal_S17e 1- 122; RIBOSOMAL_S17E 41-56;

			112,1.085; 114- 134,1.13; 31-44,1.124;	PKC_PHOSPHO_SITE 70-72; TYR_PHOSPHO_SITE 14-21; PKC_PHOSPHO_SITE 30-32; CK2_PHOSPHO_SITE 89-92; PKC_PHOSPHO_SITE 142-144;	
DEX0449_01 0.orf.1	Y	0 - 01- 166;		PKC_PHOSPHO_SITE 27-29; PKC_PHOSPHO_SITE 89-91; PKC_PHOSPHO_SITE 62-64; PKC_PHOSPHO_SITE 161-163; TYR_PHOSPHO_SITE 33-40; PKC_PHOSPHO_SITE 49-51; CK2_PHOSPHO_SITE 108-111; MYRISTYL 131-136; PKC_PHOSPHO_SITE 132-134;	Ribosomal_S17e 20- 141; RIBOSOMAL_S17E 60-75;
DEX0449_01 1.aa.1	N	0 - 01- 254;	90- 105,1.105; 221- 232,1.115; 213- 219,1.125; 15-32,1.096; 160- 173,1.13; 239- 247,1.187; 77-87,1.09;	CK2_PHOSPHO_SITE 209-212; CAMP_PHOSPHO_SITE 7-10; CK2_PHOSPHO_SITE 35-38; CK2_PHOSPHO_SITE 233-236; ASN_GLYCOSYLATION 192-195; PKC_PHOSPHO_SITE 35-37; ASN_GLYCOSYLATION 223-226; CK2_PHOSPHO_SITE 11-14; PKC_PHOSPHO_SITE 67-69; CK2_PHOSPHO_SITE 107-110; MYRISTYL 64-69; CK2_PHOSPHO_SITE 10-13; PKC_PHOSPHO_SITE 25-27; CK2_PHOSPHO_SITE 123-126; PKC_PHOSPHO_SITE 92-94; ASN_GLYCOSYLATION 65-68; TYR PHOSPHO SITE	TFIIS 216-251; ZnF_C2C2 214-253; TFS2M 91-192; TFSII 1-254; TFIIS 214- 252;

				134-140; CK2_PHOSPHO_SITE 177-180; CK2_PHOSPHO_SITE 88-91; MYRISTYL 207-212;	
DEX0449_01 1.orf.1	Y	0 - 01- 302;	208- 221,1.13; 261- 267,1.125; 4-20,1.264; 125- 134,1.09; 138- 148,1.105; 63-76,1.096; 25-32,1.092; 269- 280,1.115; 38-45,1.077; 287- 294,1.187;	CAMP_PHOSPHO_SITE 55-58; ASN_GLYCOSYLATION 271-274; CK2_PHOSPHO_SITE 59-62; PKC_PHOSPHO_SITE 140-142; CK2_PHOSPHO_SITE 83-86; TYR_PHOSPHO_SITE 182-188; CK2_PHOSPHO_SITE 281-284; CK2_PHOSPHO_SITE 58-61; MYRISTYL 255-260; CK2_PHOSPHO_SITE 136-139; ASN_GLYCOSYLATION 240-243; CK2_PHOSPHO_SITE 171-174; PKC_PHOSPHO_SITE 44-46; CK2_PHOSPHO_SITE 225-228; PKC_PHOSPHO_SITE 115-117; CK2_PHOSPHO_SITE 257-260; PKC_PHOSPHO_SITE 83-85; MYRISTYL 112-117; ASN_GLYCOSYLATION 113-116; PKC_PHOSPHO_SITE 73-75; MYRISTYL 48-53; CK2_PHOSPHO_SITE 155-158;	TFSII 2-302; TFIIS 264-299; TFS2M 139- 240; TFS2N 12-80; ZnF_C2C2 262-301; TFIIS 262-300;
DEX0449_01 2.aa.1	N	0 - 01- 225;	92-98,1.091; 159- 168,1.118; 34-43,1.152; 132- 138,1.052; 113- 125,1.136; 63-71,1.115; 4-26,1.152; 73-89,1.107; 100-	MYRISTYL 110-115; CK2_PHOSPHO_SITE 114-117; CK2_PHOSPHO_SITE 3-6; CK2_PHOSPHO_SITE 182-185; PKC_PHOSPHO_SITE 63-65; MYRISTYL 39-44; PKC_PHOSPHO_SITE 48-50;	HAD-SF-IB 16-190; serB 1-220; Hydrolase 14-200;

			110,1.184; 186- 195,1.089; 54-61,1.121; 213- 222,1.101;		
DEX0449_01 3.aa.1	N	0 - 01- 249;	33-43,1.16; 210- 218,1.114; 198- 208,1.067; 133- 155,1.217; 163- 191,1.174; 74-80,1.077; 113- 123,1.179; 232- 246,1.117; 83- 106,1.227; 54-71,1.119; 21-28,1.053; 125- 131,1.082;	PKC_PHOSPHO_SITE 16-18; CK2_PHOSPHO_SITE 216-219; MYRISTYL 41-46; PKC_PHOSPHO_SITE 204-206; MYRISTYL 212-217; PKC_PHOSPHO_SITE 15-17; MYRISTYL 33-38; CK2_PHOSPHO_SITE 47-50; TYR_PHOSPHO_SITE 61-68; CK2_PHOSPHO_SITE 162-165; MYRISTYL 138-143;	
DEX0449_01 4.aa.1	N	0 - 01- 113;	31-45,1.147; 57-80,1.147;	PKC_PHOSPHO_SITE 23-25; CK2_PHOSPHO_SITE 107-110; PKC_PHOSPHO_SITE 17-19; CK2_PHOSPHO_SITE 98-101; PKC_PHOSPHO_SITE 8-10; CK2_PHOSPHO_SITE 54-57; ASN_GLYCOSYLATION 52-55; CK2_PHOSPHO_SITE 17-20; ASN_GLYCOSYLATION 14-17;	RCC1_2 66-76;
DEX0449_01 5.aa.1	N	0 - 01- 102;	83-89,1.072; 17-38,1.144; 43-51,1.06; 65-78,1.148; 91-99,1.076;	PKC_PHOSPHO_SITE 59-61;	
DEX0449_01 6.aa.1	N	0 - 01- 412;	362- 391,1.147; 263- 294,1.164; 297- 318,1.186; 351- 359,1.147; 329-	MYRISTYL 185-190; MYRISTYL 226-231; CK2_PHOSPHO_SITE 134-137; PKC_PHOSPHO_SITE 398-400; MYRISTYL 38-43; MYRISTYL 383-388; CK2 PHOSPHO SITE	CYTOCHROME_P450 348-357; PRICHEXTENSIN 46-67; EP450I 270-288; P450 355-366; PRICHEXTENSIN 204- 229; EP450I 345- 355; EP450I 173- 190; P450 271-282;

			349,1.091; 403- 409,1.179; 44-70,1.113; 245- 254,1.08; 191- 201,1.112; 233- 243,1.061; 84- 125,1.211; 171- 182,1.135; 23-36,1.193; 5-21,1.155; 127- 159,1.183;	290-293; AMIDATION 350- 353; CK2_PHOSPHO_SITE 151-154; PKC_PHOSPHO_SITE 257-259; MYRISTYL 42-47; MYRISTYL 222-227; PKC_PHOSPHO_SITE 2-4;	EP450I 193-219; P450 184-201; p450 3-408; EP450I 355- 378; EP450I 310- 334; P450 346-355;
DEX0449_01 6.orf.1	N	0 - o1- 203;	36-45,1.08; 194- 200,1.179; 22-34,1.085; 88- 109,1.186; 8-14,1.041; 153- 182,1.147; 142- 150,1.147; 120- 140,1.091; 54-85,1.164;	PKC_PHOSPHO_SITE 48-50; PKC_PHOSPHO_SITE 15-17; AMIDATION 141-144; CK2_PHOSPHO_SITE 81-84; CK2_PHOSPHO_SITE 15-18; MYRISTYL 174-179; PKC_PHOSPHO_SITE 189-191;	EP450I 18-44; P450 27-40; BP450 146- 157; P450 146-157; EP450I 61-79; BP450 137-146; P450 62- 73; CYTOCHROME_P450 139-148; EP450IV 57-73; EP450IV 146- 164; EP450I 146- 169; EP450IV 130- 146; EP450I 136- 146; EP450I 101- 125; BP450 79-106; EP450IV 106-124; P450 137-146; BP450 62-73;
DEX0449_01 7.aa.1	Y	0 - i1- 116;	74-87,1.16; 99- 108,1.095; 4-31,1.198; 48-71,1.145;	ASN_GLYCOSYLATION 94-97; ASN_GLYCOSYLATION 64-67; PKC_PHOSPHO_SITE 98-100;	
DEX0449_01 7.orf.1	N	0 - o1- 141;	126- 138,1.16; 25-30,1.033; 100- 123,1.145; 38-44,1.04; 73-80,1.087; 87-93,1.083; 11-21,1.245; 58-68,1.175;	PKC_PHOSPHO_SITE 43-45; RGD 58-60; ASN_GLYCOSYLATION 116-119; AMIDATION 21-24; PKC_PHOSPHO_SITE 31-33; MYRISTYL 92-97; MYRISTYL 25-30; MYRISTYL 21-26;	
DEX0449_01 8.aa.1	N	0 - o1- 426;	5-43,1.154; 106- 116,1.117; 90-95,1.044; 356- 363,1.107; 224- 231.1.101;	CAMP_PHOSPHO_SITE 329-332; CK2_PHOSPHO_SITE 377-380; PKC_PHOSPHO_SITE 229-231; PKC_PHOSPHO_SITE 328-330;	ZINC_FINGER_C2H2_1 74-96;

			322- 328,1.077; 70-80,1.145; 293- 299,1.08; 205- 221,1.197; 172- 187,1.186; 192- 199,1.115; 271- 282,1.08; 154- 169,1.132; 401- 418,1.128; 59-65,1.037; 45-50,1.049; 333- 351,1.138;	CK2_PHOSPHO_SITE 135-138; CK2_PHOSPHO_SITE 75-78; TYR_PHOSPHO_SITE 264-272; CK2_PHOSPHO_SITE 120-123; CK2_PHOSPHO_SITE 137-140; PKC_PHOSPHO_SITE 368-370; CAMP_PHOSPHO_SITE 241-244; CK2_PHOSPHO_SITE 128-131; MYRISTYL 129-134; RGD 316- 318; PKC_PHOSPHO_SITE 69-71; CK2_PHOSPHO_SITE 111-114; PKC_PHOSPHO_SITE 301-303; PKC_PHOSPHO_SITE 398-400; MYRISTYL 243-248; CK2_PHOSPHO_SITE 332-335; PKC_PHOSPHO_SITE 237-239; MYRISTYL 247-252; CAMP_PHOSPHO_SITE 374-377; CK2_PHOSPHO_SITE 15-18; CK2_PHOSPHO_SITE 368-371;	
DEX0449_01 8.aa.4	N	0 - 01- 370;	237- 243,1.08; 215- 226,1.08; 345- 362,1.128; 168- 175,1.101; 266- 272,1.077; 98- 113,1.132; 116- 131,1.186; 50-60,1.117; 5-39,1.125; 149- 165,1.197; 136- 143,1.115; 300-	TYR_PHOSPHO_SITE 208-216; CK2_PHOSPHO_SITE 276-279; CK2_PHOSPHO_SITE 72-75; CK2_PHOSPHO_SITE 64-67; PKC_PHOSPHO_SITE 342-344; CAMP_PHOSPHO_SITE 185-188; CK2_PHOSPHO_SITE 79-82; PKC_PHOSPHO_SITE 181-183; CK2_PHOSPHO_SITE 55-58; PKC_PHOSPHO_SITE 312-314; PKC PHOSPHO SITE	

			307,1.107; 277- 295,1.138;	30-32; CK2_PHOSPHO_SITE 81-84; MYRISTYL 73-78; CK2_PHOSPHO_SITE 30-33; MYRISTYL 191-196; PKC_PHOSPHO_SITE 245-247; CK2_PHOSPHO_SITE 321-324; CAMP_PHOSPHO_SITE 273-276; CK2_PHOSPHO_SITE 312-315; PKC_PHOSPHO_SITE 272-274; RGD 260- 262; CK2_PHOSPHO_SITE 15-18; MYRISTYL 187-192; PKC_PHOSPHO_SITE 173-175; CAMP_PHOSPHO_SITE 318-321;	
DEX0449_01 8.aa.5	N	0 - 01- 208;	70-80,1.145; 45-50,1.049; 59-65,1.037; 90-95,1.044; 154- 169,1.132; 172- 205,1.186; 106- 116,1.117; 5-43,1.154;	CK2_PHOSPHO_SITE 111-114; CK2_PHOSPHO_SITE 120-123; PKC_PHOSPHO_SITE 69-71; CK2_PHOSPHO_SITE 135-138; MYRISTYL 129-134; CK2_PHOSPHO_SITE 75-78; CK2_PHOSPHO_SITE 128-131; CK2_PHOSPHO_SITE 15-18; CK2_PHOSPHO_SITE 137-140;	ZINC_FINGER_C2H2_1 74-96;
DEX0449_01 8.orf.7	N	0 - 01- 209;	115- 131,1.129; 17-39,1.155; 57-80,1.144; 192-203,1.2; 181- 189,1.123; 100- 106,1.067; 140- 178,1.174; 4-13,1.164;	MYRISTYL 98-103; CK2_PHOSPHO_SITE 50-53; CK2_PHOSPHO_SITE 107-110; CK2_PHOSPHO_SITE 162-165; MYRISTYL 203-208; CK2_PHOSPHO_SITE 48-51;	ank 179-209; ANK_REPEAT 179-209; ANKYRIN 192-204; ANK_REP_REGION 179- 209; ANKYRIN 180- 192; ANK 138-168; ANK 179-208; ank 138-171;
DEX0449_01 8.aa.7	N	0 - 01- 125;	70-80,1.145; 90-95,1.044; 45-50,1.049; 5-43,1.154; 59-65,1.037;	CK2_PHOSPHO_SITE 121-124; CK2_PHOSPHO_SITE 111-114; CK2 PHOSPHO SITE	ZINC_FINGER_C2H2_1 74-96;

			106- 116,1.117;	15-18; CK2_PHOSPHO_SITE 120-123; PKC_PHOSPHO_SITE 69-71; CK2_PHOSPHO_SITE 75-78;	
DEX0449_01 9.orf.1	N	0 - 01- 191;	45-58,1.139; 90- 111,1.126; 63-86,1.209; 178- 184,1.058; 115- 135,1.185; 10-37,1.175; 140- 154,1.113;	PKC_PHOSPHO_SITE 165-167; CK2_PHOSPHO_SITE 38-41; PKC_PHOSPHO_SITE 184-186; PKC_PHOSPHO_SITE 183-185; PKC_PHOSPHO_SITE 57-59; PKC_PHOSPHO_SITE 89-91; MYRISTYL 15-20; MYRISTYL 112-117; PKC_PHOSPHO_SITE 84-86;	
DEX0449_01 9.aa.1	Y	0 - 01- 130;	101- 115,1.14; 117- 127,1.111; 37-59,1.165; 13-30,1.165; 66-71,1.085; 4-11,1.144; 75-81,1.067;	CK2_PHOSPHO_SITE 79-82; AMIDATION 32-35; ASN_GLYCOSYLATION 63-66; MYRISTYL 38-43; ASN_GLYCOSYLATION 75-78; MYRISTYL 7-12; PKC_PHOSPHO_SITE 59-61; MYRISTYL 2-7;	
DEX0449_02 0.orf.1	N	0 - 01- 340;	216- 223,1.099; 183- 191,1.146; 127- 173,1.255; 269- 275,1.09; 112- 123,1.195; 315- 321,1.08; 79- 102,1.145; 226- 252,1.213; 286- 302,1.253; 23-29,1.075; 194- 213,1.151; 259- 264,1.066; 326-	CK2_PHOSPHO_SITE 223-226; MYRISTYL 285-290; PKC_PHOSPHO_SITE 267-269; TYR_PHOSPHO_SITE 322-329; PKC_PHOSPHO_SITE 17-19; CAMP_PHOSPHO_SITE 18-21; MYRISTYL 123-128; CAMP_PHOSPHO_SITE 19-22; CK2_PHOSPHO_SITE 255-258; MYRISTYL 171-176; PKC_PHOSPHO_SITE 21-23; CK2_PHOSPHO_SITE 248-251;	aminotran_4 26-325; sp_Q9BTB6_Q9BTB6_HU MAN 192-296; AA_TRANSFER_CLASS_4 212-246;

			332,1.042; 62-68,1.078;		
DEX0449_02 0.aa.1	N	0 - 01- 306;	281- 287,1.08; 28-34,1.078; 235- 241,1.09; 78-89,1.195; 192- 218,1.213; 225- 230,1.066; 93- 139,1.255; 45-68,1.145; 292- 298,1.042; 149- 157,1.146; 160- 179,1.151; 182- 189,1.099; 252- 268,1.253;	CK2_PHOSPHO_SITE 189-192; PKC_PHOSPHO_SITE 233-235; TYR_PHOSPHO_SITE 288-295; MYRISTYL 251-256; CK2_PHOSPHO_SITE 221-224; MYRISTYL 137-142; CK2_PHOSPHO_SITE 214-217; MYRISTYL 89-94;	sp_Q9BTB6_Q9BTB6_HU MAN 158-262; aminotran_4 1-291; AA_TRANSFER_CLASS_4 178-212;
DEX0449_02 0.orf.2	N	0 - 01- 485;	224- 247,1.145; 328- 336,1.146; 145- 152,1.082; 339- 358,1.151; 120- 131,1.063; 207- 213,1.078; 371- 397,1.213; 272- 318,1.255; 361- 368,1.099; 176- 195,1.149; 257- 268,1.195; 404- 412,1.107; 464- 476,1.259; 104- 118,1.181; 427- 449,1.147; 84- 102,1.102; 26-33,1.101; 4-18,1.182;	MYRISTYL 46-51; MYRISTYL 460-465; PKC_PHOSPHO_SITE 481-483; ASN_GLYCOSYLATION 178-181; MYRISTYL 47-52; CK2_PHOSPHO_SITE 393-396; MYRISTYL 55-60; PKC_PHOSPHO_SITE 135-137; MYRISTYL 316-321; MYRISTYL 43-48; PKC_PHOSPHO_SITE 124-126; MYRISTYL 77-82; PKC_PHOSPHO_SITE 56-58; PKC_PHOSPHO_SITE 450-452; MYRISTYL 100-105; CK2_PHOSPHO_SITE 450-453; CK2_PHOSPHO_SITE 368-371; MYRISTYL 268-273; ASN_GLYCOSYLATION 33-36; MYRISTYL 76-81; CK2_PHOSPHO_SITE 153-156;	aminotran_4 170- 453; sp_Q9BTB6_Q9BTB6_HU MAN 337-401; AA_TRANSFER_CLASS_4 357-391;

DEX0449_02 0.aa.2	N	0 - 01- 465;	11-25,1.181; 394- 400,1.09; 451- 457,1.042; 52-59,1.082; 164- 175,1.195; 319- 324,1.061; 371- 389,1.133; 361- 368,1.14; 246- 265,1.151; 278- 304,1.213; 114- 120,1.078; 4-9,1.069; 179- 225,1.255; 235- 243,1.146; 411- 427,1.253; 440- 446,1.08; 83- 102,1.149; 27-38,1.063; 334- 357,1.22; 268- 275,1.099; 131- 154,1.145;	ASN_GLYCOSYLATION 85-88; MYRISTYL 7-12; TYR_PHOSPHO_SITE 447-454; CK2_PHOSPHO_SITE 314-317; MYRISTYL 175-180; CK2_PHOSPHO_SITE 275-278; CK2_PHOSPHO_SITE 60-63; CK2_PHOSPHO_SITE 300-303; PKC_PHOSPHO_SITE 31-33; MYRISTYL 333-338; MYRISTYL 323-328; MYRISTYL 410-415; PKC_PHOSPHO_SITE 42-44; MYRISTYL 223-228; PKC_PHOSPHO_SITE 392-394; CK2_PHOSPHO_SITE 307-310;	sp_Q9BTB6_Q9BTB6_HU MAN 382-421; sp_Q9BTB6_Q9BTB6_HU MAN 244-310; aminotran_4 77-450; AA_TRANSFER_CLASS_4 264-298;
DEX0449_02 1.orf.1	N	0 - 01- 74;		ASN_GLYCOSYLATION 7-10; MYRISTYL 8- 13; PKC_PHOSPHO_SITE 11-13; PKC_PHOSPHO_SITE 13-15; CK2_PHOSPHO_SITE 1-4; CAMP_PHOSPHO_SITE 17-20; MYRISTYL 32-37; ASN_GLYCOSYLATION 10-13; PKC_PHOSPHO_SITE 39-41; PKC_PHOSPHO_SITE 1-3; CK2_PHOSPHO_SITE 11-14; ASN_GLYCOSYLATION	

				9-12; MYRISTYL 3-8; PKC_PHOSPHO_SITE 12-14;	
DEX0449_02 1.aa.1	N	1 - i1- 12;tm13 - 32;o33- 63;	50-60,1.136; 4-47,1.195;		
DEX0449_02 2.aa.1	N	0 - o1- 313;	119- 137,1.178; 217- 247,1.231; 76-83,1.102; 24-30,1.066; 267- 282,1.131; 88- 116,1.315; 250- 257,1.128; 294- 310,1.104; 149- 155,1.04; 165- 210,1.137; 139- 147,1.071; 65-71,1.081; 40-59,1.204;	PKC_PHOSPHO_SITE 12-14; MYRISTYL 246-251; CK2_PHOSPHO_SITE 7-10; MYRISTYL 65-70; PKC_PHOSPHO_SITE 7-9; CK2_PHOSPHO_SITE 143-146; PKC_PHOSPHO_SITE 118-120; PKC_PHOSPHO_SITE 269-271; CK2_PHOSPHO_SITE 15-18; ASN_GLYCOSYLATION 22-25; CK2_PHOSPHO_SITE 255-258; CK2_PHOSPHO_SITE 159-162; PKC_PHOSPHO_SITE 147-149; CK2_PHOSPHO_SITE 113-116; PKC_PHOSPHO_SITE 193-195; CK2_PHOSPHO_SITE 147-150; ASN_GLYCOSYLATION 111-114; CK2_PHOSPHO_SITE 269-272; PKC_PHOSPHO_SITE 88-90; MYRISTYL 289-294; PKC_PHOSPHO_SITE 292-294; CK2_PHOSPHO_SITE 61-64;	Fe_hyd_lg_C 79-299;
DEX0449_02 2.orf.2	N	0 - o1- 395;	345- 392,1.137; 87- 118,1.108; 245- 251,1.081; 256- 263,1.102; 220-	CK2_PHOSPHO_SITE 323-326; PKC_PHOSPHO_SITE 78-80; PKC_PHOSPHO_SITE 298-300; PKC_PHOSPHO_SITE 46-48; MYRISTYL 110-115; MYRISTYL	Fe_hyd_lg_C 259- 395; CLOACIN 159- 180; CLOACIN 270- 289; PRO_RICH 18- 105;

			239,1.204; 122- 143,1.17; 319- 327,1.071; 268- 296,1.315; 39-58,1.105; 26-33,1.048; 171- 186,1.11; 204- 210,1.066; 157- 169,1.174; 60-65,1.04; 299- 317,1.178; 329- 335,1.04;	45-50; MYRISTYL 245-250; CK2_PHOSPHO_SITE 8-11; CAMP_PHOSPHO_SITE 5-8; ASN_GLYCOSYLATION 291-294; CK2_PHOSPHO_SITE 241-244; CK2_PHOSPHO_SITE 146-149; CAMP_PHOSPHO_SITE 80-83; CK2_PHOSPHO_SITE 83-86; CK2_PHOSPHO_SITE 195-198; CK2_PHOSPHO_SITE 187-190; CK2_PHOSPHO_SITE 35-38; PKC_PHOSPHO_SITE 192-194; PKC_PHOSPHO_SITE 268-270; CK2_PHOSPHO_SITE 293-296; MYRISTYL 141-146; ASN_GLYCOSYLATION 202-205; CK2_PHOSPHO_SITE 339-342; PKC_PHOSPHO_SITE 373-375; MYRISTYL 159-164; MYRISTYL 173-178; CK2_PHOSPHO_SITE 327-330; PKC_PHOSPHO_SITE 327-329; PKC_PHOSPHO_SITE 187-189;	
DEX0449_02 2.aa.2	N	0 - 01- 285;	233- 241,1.12; 193- 207,1.205; 125- 134,1.092; 166- 176,1.158; 184- 189,1.045; 46-76,1.231; 4-10,1.133; 145- 150,1.073; 79-86,1.128; 222-	CK2_PHOSPHO_SITE 84-87; TYR_PHOSPHO_SITE 227-233; PKC_PHOSPHO_SITE 260-262; MYRISTYL 31-36; PKC_PHOSPHO_SITE 148-150; MYRISTYL 75-80; PKC_PHOSPHO_SITE 98-100; MYRISTYL 205-210; MYRISTYL 10-15; CK2_PHOSPHO_SITE 35-38;	Fe_hyd_lg_C 1-214; Fe_hyd_SSU 220-275;

			227,1.076; 96- 110,1.092; 244- 253,1.118; 38-44,1.11; 263- 271,1.105;	CK2_PHOSPHO_SITE 117-120; MYRISTYL 206-211; CK2_PHOSPHO_SITE 98-101;	
DEX0449_02 3.orf.1	N	0 - ol- 170;	133- 143,1.109; 50-73,1.14; 81- 115,1.154; 36-48,1.177; 12-30,1.165; 148- 153,1.079;	PKC_PHOSPHO_SITE 161-163; MYRISTYL 143-148; PKC_PHOSPHO_SITE 118-120; PKC_PHOSPHO_SITE 119-121; MYRISTYL 127-132; CAMP_PHOSPHO_SITE 129-132; PKC_PHOSPHO_SITE 128-130; PKC_PHOSPHO_SITE 50-52;	
DEX0449_02 3.aa.1	N	0 - il- 159;	25-37,1.177; 122- 132,1.109; 39-62,1.14; 137- 142,1.079; 70- 104,1.154; 4-19,1.165;	PKC_PHOSPHO_SITE 117-119; PKC_PHOSPHO_SITE 107-109; PKC_PHOSPHO_SITE 150-152; MYRISTYL 132-137; CAMP_PHOSPHO_SITE 118-121; PKC_PHOSPHO_SITE 39-41; PKC_PHOSPHO_SITE 108-110; MYRISTYL 116-121;	
DEX0449_02 3.orf.2	N	0 - ol- 170;	81- 115,1.154; 133- 143,1.109; 12-30,1.165; 148- 153,1.079; 50-73,1.14; 36-48,1.177;	PKC_PHOSPHO_SITE 161-163; PKC_PHOSPHO_SITE 50-52; MYRISTYL 143-148; PKC_PHOSPHO_SITE 119-121; CAMP_PHOSPHO_SITE 129-132; PKC_PHOSPHO_SITE 128-130; PKC_PHOSPHO_SITE 118-120; MYRISTYL 127-132;	
DEX0449_02 4.orf.1	N	0 - ol- 255;	136- 145,1.106; 228- 249,1.154; 182- 192,1.188; 4-20,1.125; 204-	PKC_PHOSPHO_SITE 150-152; MYRISTYL 228-233; MYRISTYL 225-230; CK2_PHOSPHO_SITE 241-244; CK2_PHOSPHO_SITE 250-253;	

			210,1.043; 213- 224,1.184; 84-90,1.092; 28-78,1.171; 117- 123,1.077; 106- 115,1.076;	PKC_PHOSPHO_SITE 24-26; MYRISTYL 234-239; PKC_PHOSPHO_SITE 92-94; MYRISTYL 170-175; PKC_PHOSPHO_SITE 81-83; MYRISTYL 55-60; CK2_PHOSPHO_SITE 200-203; PKC_PHOSPHO_SITE 166-168; CK2_PHOSPHO_SITE 92-95; CK2_PHOSPHO_SITE 24-27; MYRISTYL 210-215; MYRISTYL 230-235; ASN_GLYCOSYLATION 98-101;	
DEX0449_02 4.aa.1	N	0 - ol- 147;	35-44,1.174; 109- 115,1.09; 97- 105,1.083; 121- 134,1.19; 5- 23,1.121; 87-94,1.123; 52-74,1.216;	MYRISTYL 65-70; CK2_PHOSPHO_SITE 93-96; PKC_PHOSPHO_SITE 54-56; CK2_PHOSPHO_SITE 80-83; PKC_PHOSPHO_SITE 23-25; CK2_PHOSPHO_SITE 135-138; PKC_PHOSPHO_SITE 108-110; PKC_PHOSPHO_SITE 45-47;	CSD 62-129; CSP 64- 129; COLD SHOCK 65- 80; COLD SHOCK 86- 95;
DEX0449_02 4.orf.2	N	0 - ol- 255;	213- 224,1.184; 182- 192,1.188; 84-90,1.092; 117- 123,1.077; 28-78,1.171; 136- 145,1.106; 4-20,1.125; 204- 210,1.043; 228- 249,1.154; 106- 115,1.076;	CK2_PHOSPHO_SITE 200-203; PKC_PHOSPHO_SITE 24-26; CK2_PHOSPHO_SITE 241-244; PKC_PHOSPHO_SITE 150-152; CK2_PHOSPHO_SITE 24-27; MYRISTYL 225-230; MYRISTYL 230-235; MYRISTYL 55-60; MYRISTYL 234-239; MYRISTYL 170-175; PKC_PHOSPHO_SITE 92-94; PKC_PHOSPHO_SITE 166-168; CK2_PHOSPHO_SITE 250-253; MYRISTYL 228-233; ASN_GLYCOSYLATION	

				98-101; PKC_PHOSPHO_SITE 81-83; CK2_PHOSPHO_SITE 92-95; MYRISTYL 210-215;	
DEX0449_02 4.orf.3	N	0 - 01- 243;	94- 103,1.076; 192- 198,1.043; 170- 180,1.188; 216- 237,1.154; 16-66,1.171; 72-78,1.092; 201- 212,1.184; 124- 133,1.106; 105- 111,1.077;	MYRISTYL 9-14; CK2_PHOSPHO_SITE 238-241; CK2_PHOSPHO_SITE 229-232; MYRISTYL 216-221; CK2_PHOSPHO_SITE 188-191; MYRISTYL 222-227; MYRISTYL 218-223; MYRISTYL 8-13; PKC_PHOSPHO_SITE 69-71; MYRISTYL 213-218; PKC_PHOSPHO_SITE 154-156; MYRISTYL 43-48; PKC_PHOSPHO_SITE 138-140; MYRISTYL 158-163; PKC_PHOSPHO_SITE 80-82; ASN_GLYCOSYLATION 86-89; CK2_PHOSPHO_SITE 80-83; PKC_PHOSPHO_SITE 3-5; MYRISTYL 198-203;	
DEX0449_02 4.aa.3	N	0 - 01- 147;	5-23,1.121; 52-74,1.216; 121- 134,1.19; 109- 115,1.09; 35-44,1.174; 97- 105,1.083; 87-94,1.123;	PKC_PHOSPHO_SITE 54-56; CK2_PHOSPHO_SITE 80-83; PKC_PHOSPHO_SITE 45-47; CK2_PHOSPHO_SITE 135-138; PKC_PHOSPHO_SITE 23-25; CK2_PHOSPHO_SITE 93-96; MYRISTYL 65-70; PKC_PHOSPHO_SITE 108-110;	CSP 64-129; COLDSHOCK 65-80; COLDSHOCK 86-95; CSD 62-129;
DEX0449_02 4.orf.4	N	0 - 01- 202;	129- 139,1.188; 160- 171,1.184; 23-37,1.092; 151- 157,1.043; 53-62,1.076; 175-	MYRISTYL 181-186; MYRISTYL 175-180; ASN_GLYCOSYLATION 45-48; PKC_PHOSPHO_SITE 29-31; MYRISTYL 172-177; CK2_PHOSPHO_SITE 188-191;	

			196,1.154; 64-70,1.077; 11-17,1.151; 83-92,1.106;	PKC_PHOSPHO_SITE 39-41; MYRISTYL 117-122; MYRISTYL 2-7; PKC_PHOSPHO_SITE 113-115; CK2_PHOSPHO_SITE 39-42; CK2_PHOSPHO_SITE 197-200; CK2_PHOSPHO_SITE 147-150; MYRISTYL 157-162; PKC_PHOSPHO_SITE 97-99; MYRISTYL 177-182;	
DEX0449_02 4.aa.4	N	0 - 01- 138;	108- 119,1.184; 99- 105,1.043; 123- 135,1.126; 4-10,1.076; 12-18,1.077; 31-40,1.106; 77-87,1.188;	CK2_PHOSPHO_SITE 95-98; MYRISTYL 105-110; PKC_PHOSPHO_SITE 45-47; MYRISTYL 123-128; MYRISTYL 120-125; PKC_PHOSPHO_SITE 61-63; MYRISTYL 65-70;	
DEX0449_02 4.aa.5	N	0 - 01- 185;	124- 134,1.188; 78-87,1.106; 155- 166,1.184; 59-65,1.077; 48-57,1.076; 170- 182,1.126; 146- 152,1.043; 5-23,1.121;	CK2_PHOSPHO_SITE 142-145; MYRISTYL 170-175; CK2_PHOSPHO_SITE 34-37; PKC_PHOSPHO_SITE 92-94; PKC_PHOSPHO_SITE 34-36; MYRISTYL 152-157; PKC_PHOSPHO_SITE 23-25; PKC_PHOSPHO_SITE 108-110; MYRISTYL 112-117; MYRISTYL 167-172; ASN_GLYCOSYLATION 40-43;	
DEX0449_02 4.orf.5	N	0 - 01- 265;	214- 220,1.043; 4-14,1.143; 223- 234,1.184; 146- 155,1.106; 127- 133,1.077; 116- 125,1.076; 73-91,1.121; 192- 202,1.188; 57-65,1.074;	PKC_PHOSPHO_SITE 102-104; PKC_PHOSPHO_SITE 176-178; MYRISTYL 27-32; PKC_PHOSPHO_SITE 91-93; MYRISTYL 235-240; MYRISTYL 31-36; PKC_PHOSPHO_SITE 51-53; MYRISTYL 238-243; PKC_PHOSPHO_SITE 160-162; CK2 PHOSPHO SITE	

			238- 259,1.154; 18-27,1.167;	260-263; MYRISTYL 180-185; MYRISTYL 240-245; CK2_PHOSPHO_SITE 102-105; MYRISTYL 42-47; CK2_PHOSPHO_SITE 210-213; MYRISTYL 220-225; CK2_PHOSPHO_SITE 251-254; MYRISTYL 244-249; PKC_PHOSPHO_SITE 15-17; ASN_GLYCOSYLATION 108-111;	
DEX0449_02 4.aa.6	Y	0 - o1- 201;	42-56,1.099; 33-39,1.074; 58- 181,1.229; 14-31,1.197;	MYRISTYL 95-100; PKC_PHOSPHO_SITE 30-32; CK2_PHOSPHO_SITE 127-130; MYRISTYL 143-148; CK2_PHOSPHO_SITE 10-13; PKC_PHOSPHO_SITE 163-165; PKC_PHOSPHO_SITE 10-12; CK2_PHOSPHO_SITE 74-77; MYRISTYL 67-72; PKC_PHOSPHO_SITE 181-183;	ER_TARGET 198-201;
DEX0449_02 5.aa.1	N	0 - o1- 159;	149- 155,1.129; 17-32,1.132; 109- 136,1.177; 74-88,1.162; 49-67,1.112; 99- 107,1.157; 34-43,1.178;	CK2_PHOSPHO_SITE 138-141; CK2_PHOSPHO_SITE 41-44; PKC_PHOSPHO_SITE 85-87; MYRISTYL 144-149; CK2_PHOSPHO_SITE 91-94; PKC_PHOSPHO_SITE 52-54; CK2_PHOSPHO_SITE 108-111; PKC_PHOSPHO_SITE 29-31; CK2_PHOSPHO_SITE 143-146;	
DEX0449_02 5.aa.2	N	1 - i1- 36;tm37 - 54;o55- 128;	35-44,1.203; 4-10,1.103; 118- 124,1.129; 48-56,1.102; 68-76,1.157; 78- 105,1.177;	CK2_PHOSPHO_SITE 77-80; CK2_PHOSPHO_SITE 60-63; ASN_GLYCOSYLATION 44-47; MYRISTYL 113-118; CK2_PHOSPHO_SITE 107-110;	

				PKC_PHOSPHO_SITE 57-59; ASN_GLYCOSYLATION 12-15; CK2_PHOSPHO_SITE 112-115; PKC_PHOSPHO_SITE 13-15;	
DEX0449_02 6.aa.1	N	0 - 01- 224;	63-74, 1.101; 148- 166, 1.202; 43-49, 1.048; 186- 195, 1.118; 135- 141, 1.07; 16-28, 1.137; 124- 129, 1.03; 114- 120, 1.096; 170- 178, 1.135; 198- 205, 1.103; 90-96, 1.08; 212- 221, 1.109;	CK2_PHOSPHO_SITE 119-122; CAMP_PHOSPHO_SITE 143-146; CK2_PHOSPHO_SITE 188-191; ASN_GLYCOSYLATION 59-62; PKC_PHOSPHO_SITE 77-79; CK2_PHOSPHO_SITE 101-104; PKC_PHOSPHO_SITE 141-143; CK2_PHOSPHO_SITE 108-111; PKC_PHOSPHO_SITE 37-39; TYR_PHOSPHO_SITE 166-174; PKC_PHOSPHO_SITE 142-144;	SH2DOMAIN 113-127; SH2 113-196; SH2 111-202; PRO_RICH 19-102; SH2 113- 221; sp_Q13094_LCP2_HUMA N 113-207; SH2DOMAIN 185-199; SH2DOMAIN 133-143;
DEX0449_02 6.aa.2	Y	0 - 01- 230;	176- 184, 1.135; 154- 172, 1.202; 120- 126, 1.096; 49-55, 1.048; 218- 227, 1.109; 204- 211, 1.103; 141- 147, 1.07; 69-80, 1.101; 8-34, 1.204; 130- 135, 1.03; 192- 201, 1.118; 96-102, 1.08;	TYR_PHOSPHO_SITE 172-180; PKC_PHOSPHO_SITE 148-150; PKC_PHOSPHO_SITE 43-45; CK2_PHOSPHO_SITE 114-117; PKC_PHOSPHO_SITE 83-85; MYRISTYL 8-13; PKC_PHOSPHO_SITE 147-149; CAMP_PHOSPHO_SITE 149-152; CK2_PHOSPHO_SITE 107-110; CK2_PHOSPHO_SITE 194-197; CK2_PHOSPHO_SITE 125-128; ASN_GLYCOSYLATION 65-68;	SH2DOMAIN 119-133; PRO_RICH 25-108; sp_Q13094_LCP2_HUMA N 119-213; SH2 119- 227; SH2DOMAIN 191- 205; SH2 117-208; SH2 119-202; SH2DOMAIN 139-149;
DEX0449_02 6.orf.2	Y	0 - 01- 225;	91-97, 1.08; 136- 142, 1.07; 149- 167, 1.202; 213-	PKC_PHOSPHO_SITE 78-80; PKC_PHOSPHO_SITE 38-40; CK2_PHOSPHO_SITE 189-192;	SH2 114-197; SH2 114-222; sp_Q13094_LCP2_HUMA N 114-208; PRO_RICH 20-103; SH2 112- 203; SH2DOMAIN 186-

			222,1.109; 187- 196,1.118; 64-75,1.101; 125- 130,1.03; 171- 179,1.135; 199- 206,1.103; 44-50,1.048; 4-29,1.249; 115- 121,1.096;	CK2_PHOSPHO_SITE 102-105; PKC_PHOSPHO_SITE 142-144; CK2_PHOSPHO_SITE 120-123; ASN_GLYCOSYLATION 60-63; PKC_PHOSPHO_SITE 143-145; CK2_PHOSPHO_SITE 109-112; CAMP_PHOSPHO_SITE 144-147; TYR_PHOSPHO_SITE 167-175;	200; SH2DOMAIN 114- 128; SH2DOMAIN 134- 144;
DEX0449_02 7.aa.1	N	0 - 01- 1085;	1028- 1040,1.222; 860- 866,1.065; 684- 694,1.255; 13-34,1.134; 660- 676,1.101; 1008- 1020,1.107; 176- 185,1.108; 784- 811,1.127; 720- 736,1.173; 205- 223,1.113; 981- 994,1.16; 36-43,1.066; 426- 450,1.214; 613- 624,1.087; 957- 973,1.166; 585- 592,1.129; 311- 317,1.111; 357- 369,1.131; 286- 309,1.102; 704- 713,1.132; 117- 129,1.172; 884- 894,1.151; 747-	CK2_PHOSPHO_SITE 80-83; MYRISTYL 99-104; MYRISTYL 98-103; CK2_PHOSPHO_SITE 181-184; MYRISTYL 335-340; MYRISTYL 95-100; MYRISTYL 512-517; MYRISTYL 96-101; MYRISTYL 683-688; CK2_PHOSPHO_SITE 116-119; MYRISTYL 42-47; CK2_PHOSPHO_SITE 580-583; MYRISTYL 889-894; ASN_GLYCOSYLATION 766-769; TYR_PHOSPHO_SITE 323-330; MYRISTYL 656-661; CK2_PHOSPHO_SITE 299-302; PKC_PHOSPHO_SITE 390-392; MYRISTYL 679-684; MYRISTYL 12-17; MYRISTYL 738-743; CAMP_PHOSPHO_SITE 236-239; ASN_GLYCOSYLATION 456-459; PKC_PHOSPHO_SITE 680-682; MYRISTYL 63-68; MYRISTYL 76-81; ASN_GLYCOSYLATION 783-786; PKC_PHOSPHO_SITE 939-941; MYRISTYL 93-98; PKC PHOSPHO SITE	LON 123-368; LON_SER 978-986; ENDOLAPTASE 1005- 1024; GLY_RICH 56- 106; LON 123-368; AAA 515-786; ENDOLAPTASE 1028- 1046; ATP_GTP_A 523-530; lon 125- 1073; ENDOLAPTASE 975-994; DISEASERSIST 579- 593; DISEASERSIST 518-533; ENDOLAPTASE 523- 542; ENDOLAPTASE 888-904; AAA 518- 839;

			766,1.195; 251- 257,1.062; 639- 648,1.073; 516- 530,1.143; 131- 144,1.103; 602- 608,1.102; 264- 283,1.144; 815- 833,1.167; 738- 745,1.081; 1056- 1072,1.154; 941- 954,1.059; 925- 934,1.076; 321- 333,1.112; 480- 489,1.102; 869- 880,1.117; 146- 169,1.172; 533- 539,1.047; 499- 510,1.138; 1046- 1054,1.093; 571- 579,1.12; 462- 468,1.049; 45-53,1.059; 835- 858,1.106; 563- 569,1.04; 372- 378,1.057; 903- 909,1.107; 106- 112,1.104; 392- 406,1.113; 5-11,1.071;	1012-1014; MYRISTYL 634-639; PKC_PHOSPHO_SITE 975-977; MYRISTYL 89-94; CK2_PHOSPHO_SITE 975-978; AMIDATION 241- 244; PKC_PHOSPHO_SITE 72-74; CK2_PHOSPHO_SITE 339-342; ASN_GLYCOSYLATION 450-453; PKC_PHOSPHO_SITE 299-301; CK2_PHOSPHO_SITE 770-773; MYRISTYL 893-898; MYRISTYL 675-680; PKC_PHOSPHO_SITE 251-253; MYRISTYL 734-739; MYRISTYL 744-749; MYRISTYL 735-740; PKC_PHOSPHO_SITE 239-241; MYRISTYL 654-659; MYRISTYL 77-82; MYRISTYL 81-86; MYRISTYL 809-814; MYRISTYL 526-531; ASN_GLYCOSYLATION 174-177; PKC_PHOSPHO_SITE 909-911; MYRISTYL 446-451; MYRISTYL 637-642; MYRISTYL 567-572;	
DEX0449_02 7.orf.1	N	0 - ol- 668;	444- 458,1.113; 615- 621.1.04:	CK2_PHOSPHO_SITE 132-135; MYRISTYL 387-392; MYRISTYL 145-150:	GLY_RICH 108-158; DISEASERSIST 570- 585; DISEASERSIST 631-645: ATP GTP A

			41-48,1.124; 183- 196,1.103; 532- 541,1.102; 585- 591,1.047; 478- 502,1.214; 623- 631,1.12; 363- 369,1.111; 424- 430,1.057; 88-95,1.066; 257- 275,1.113; 169- 181,1.172; 551- 562,1.138; 198- 221,1.172; 10-16,1.06; 57-63,1.071; 514- 520,1.049; 316- 335,1.144; 65-86,1.134; 158- 164,1.104; 30-36,1.085; 228- 237,1.108; 637- 644,1.129; 654- 660,1.102; 409- 421,1.131; 303- 309,1.062; 338- 361,1.102; 568- 582,1.143; 373- 385,1.112; 97- 105,1.059;	AMIDATION 293- 296; ASN_GLYCOSYLATION 226-229; MYRISTYL 39-44; MYRISTYL 141-146; PKC_PHOSPHO_SITE 303-305; CK2_PHOSPHO_SITE 632-635; CK2_PHOSPHO_SITE 233-236; MYRISTYL 133-138; CK2_PHOSPHO_SITE 391-394; CK2_PHOSPHO_SITE 168-171; MYRISTYL 128-133; MYRISTYL 147-152; MYRISTYL 150-155; PKC_PHOSPHO_SITE 291-293; CK2_PHOSPHO_SITE 23-26; MYRISTYL 148-153; PKC_PHOSPHO_SITE 12-14; CK2_PHOSPHO_SITE 38-41; MYRISTYL 64-69; PKC_PHOSPHO_SITE 124-126; MYRISTYL 578-583; ASN_GLYCOSYLATION 508-511; PKC_PHOSPHO_SITE 442-444; MYRISTYL 115-120; CAMP_PHOSPHO_SITE 288-291; MYRISTYL 564-569; MYRISTYL 498-503; MYRISTYL 619-624; CK2_PHOSPHO_SITE 351-354; PKC_PHOSPHO_SITE 351-353; CAMP_PHOSPHO_SITE 8-11; MYRISTYL 151-156; MYRISTYL 129-134; TYR_PHOSPHO_SITE 375-382; ASN_GLYCOSYLATION 502-505; PKC_PHOSPHO_SITE 6-8; MYRISTYL 94- 99;	575-582; AAA 567- 665; LON 175-420; LON 175-420;
DEX0449 02	N	0 - 01-	345-	TYR PHOSPHO SITE	lon 3-759: AAA 330-

7.aa.2	771;	351,1.047; 470- 497,1.127; 238- 262,1.214; 204- 218,1.113; 611- 620,1.076; 501- 519,1.167; 133- 145,1.112; 570- 580,1.151; 643- 659,1.166; 714- 726,1.222; 667- 680,1.16; 627- 640,1.059; 521- 544,1.106; 742- 758,1.154; 589- 595,1.107; 79-91,1.062; 732- 740,1.093; 6-19,1.089; 169- 181,1.131; 123- 129,1.111; 184- 190,1.057; 397- 404,1.129; 311- 322,1.138; 32-57,1.174; 432- 452,1.18; 546- 552,1.065; 328- 342,1.143; 416- 425,1.132; 95- 121,1.152; 694- 706,1.107; 292- 301,1.102; 274-	135-142; ASN_GLYCOSYLATION 469-472; MYRISTYL 64-69; MYRISTYL 579-584; MYRISTYL 575-580; CK2_PHOSPHO_SITE 661-664; PKC_PHOSPHO_SITE 698-700; ASN_GLYCOSYLATION 452-455; CK2_PHOSPHO_SITE 151-154; ASN_GLYCOSYLATION 268-271; ASN_GLYCOSYLATION 262-265; CK2_PHOSPHO_SITE 111-114; CK2_PHOSPHO_SITE 456-459; MYRISTYL 258-263; PKC_PHOSPHO_SITE 661-663; PKC_PHOSPHO_SITE 111-113; MYRISTYL 68-73; PKC_PHOSPHO_SITE 202-204; PKC_PHOSPHO_SITE 595-597; CK2_PHOSPHO_SITE 392-395; MYRISTYL 102-107; MYRISTYL 379-384; PKC_PHOSPHO_SITE 625-627; MYRISTYL 338-343; MYRISTYL 324-329; MYRISTYL 147-152; CAMP_PHOSPHO_SITE 4-7; MYRISTYL 73- 78; MYRISTYL 495- 500;	525; ENDOLAPTASE 574-590; AAA 327- 472; LON_SER 664- 672; DISEASERSIST 391-405; LON 2-180; DISEASERSIST 330- 345; ENDOLAPTASE 335-354; ENDOLAPTASE 714- 732; ENDOLAPTASE 691-710; LON 5-180; CLPPROTEASEA 441- 455; ATP_GTP_A 335- 342; CLPPROTEASEA 331-349; ENDOLAPTASE 661- 680;
--------	------	--	---	---

			280,1.049; 555- 566,1.117; 383- 391,1.12; 375- 381,1.04;		
DEX0449_02 7.aa.3	N	0 - o1- 848;	321- 333,1.112; 205- 223,1.113; 499- 510,1.138; 131- 144,1.103; 36-43,1.066; 286- 309,1.102; 533- 539,1.047; 585- 592,1.129; 146- 169,1.172; 426- 450,1.214; 117- 129,1.172; 813- 819,1.153; 176- 185,1.108; 480- 489,1.102; 658- 685,1.127; 839- 845,1.126; 743- 754,1.117; 251- 257,1.062; 604- 613,1.132; 734- 740,1.065; 392- 406,1.113; 784- 790,1.079; 13-34,1.134; 516- 530,1.143; 264- 283,1.144; 758- 768,1.151; 709- 732.1.106;	CK2_PHOSPHO_SITE 181-184; ASN_GLYCOSYLATION 174-177; MYRISTYL 810-815; MYRISTYL 526-531; MYRISTYL 93-98; MYRISTYL 767-772; MYRISTYL 76-81; CK2_PHOSPHO_SITE 116-119; ASN_GLYCOSYLATION 657-660; MYRISTYL 512-517; AMIDATION 241- 244; CK2_PHOSPHO_SITE 80-83; ASN_GLYCOSYLATION 640-643; MYRISTYL 89-94; CAMP_PHOSPHO_SITE 236-239; MYRISTYL 446-451; MYRISTYL 763-768; MYRISTYL 98-103; MYRISTYL 567-572; PKC_PHOSPHO_SITE 239-241; MYRISTYL 824-829; CK2_PHOSPHO_SITE 299-302; MYRISTYL 99-104; MYRISTYL 12-17; MYRISTYL 683-688; MYRISTYL 42-47; MYRISTYL 95-100; ASN_GLYCOSYLATION 450-453; MYRISTYL 77-82; PKC_PHOSPHO_SITE 72-74; MYRISTYL 81-86; MYRISTYL 96-101; PKC_PHOSPHO_SITE 299-301; MYRISTYL 63-68; TYR_PHOSPHO_SITE 323-330; PKC_PHOSPHO_SITE 251-253; PKC PHOSPHO SITE	lon 125-842; AAA 515-660; DISEASERSIST 579- 593; LON 123-368; CLPPROTEASEA 629- 643; LIPOCALIN 803- 815; AAA 518-713; ENDOLAPTASE 523- 542; ENDOLAPTASE 762-778; CLPPROTEASEA 519- 537; ATP_GTP_A 523- 530; GLY_RICH 56- 106; DISEASERSIST 518-533; LON 123- 368;

			563- 569,1.04; 357- 369,1.131; 462- 468,1.049; 311- 317,1.111; 571- 579,1.12; 45-53,1.059; 794- 809,1.217; 620- 640,1.18; 372- 378,1.057; 106- 112,1.104; 689- 707,1.167; 5-11,1.071;	792-794; AMIDATION 778- 781; CK2_PHOSPHO_SITE 580-583; CK2_PHOSPHO_SITE 339-342; ASN_GLYCOSYLATION 456-459; CK2_PHOSPHO_SITE 644-647; MYRISTYL 335-340; PKC_PHOSPHO_SITE 390-392;	
DEX0449_02 8.aa.1	N	0 - o1- 124;	4-14,1.157; 110- 116,1.024; 37-45,1.061; 48-53,1.048; 62-79,1.175;	MYRISTYL 115-120; CK2_PHOSPHO_SITE 15-18; TYR_PHOSPHO_SITE 58-66; CK2_PHOSPHO_SITE 99-102; PKC_PHOSPHO_SITE 37-39; TYR_PHOSPHO_SITE 45-53; ASN_GLYCOSYLATION 90-93; CK2_PHOSPHO_SITE 30-33;	
DEX0449_02 9.aa.1	N	1 - i1- 16;tm17 - 34;o35- 296;	11-36,1.154; 181- 194,1.132; 261- 267,1.144; 213- 221,1.069; 95- 119,1.183; 60-70,1.115; 247- 254,1.114; 148- 161,1.085; 125- 139,1.129; 167- 177,1.148;	CK2_PHOSPHO_SITE 242-245; PKC_PHOSPHO_SITE 3-5; MYRISTYL 258-263; CK2_PHOSPHO_SITE 40-43; PKC_PHOSPHO_SITE 219-221; PKC_PHOSPHO_SITE 146-148; MYRISTYL 155-160; ASN_GLYCOSYLATION 239-242; PKC_PHOSPHO_SITE 179-181; CK2_PHOSPHO_SITE 195-198; MYRISTYL 22-27; CK2_PHOSPHO_SITE 179-182; CAMP PHOSPHO SITE	KH 51-120; KH_TYPE_1_1 52-115; KH 56-103; KH 128- 178; KH 123-195; KH_TYPE_1_2 124- 190;

				79-82;	
DEX0449_03 0.aa.1	N	1 - i1- 164;tml 65- 187;o18 8-195;	11-17,1.104; 83- 101,1.178; 140- 151,1.186; 24-32,1.128; 42-78,1.18; 114- 126,1.137; 165- 192,1.253; 105- 112,1.023;	PKC_PHOSPHO_SITE 161-163; PKC_PHOSPHO_SITE 99-101; MYRISTYL 181-186; ASN_GLYCOSYLATION 21-24; PKC_PHOSPHO_SITE 68-70; MYRISTYL 185-190; MYRISTYL 158-163; AMIDATION 191- 194; ASN_GLYCOSYLATION 129-132; MYRISTYL 171-176; CK2_PHOSPHO_SITE 131-134; ASN_GLYCOSYLATION 93-96;	MHCCLASSI 136-154; MHCCLASSI 65-81;
DEX0449_03 0.orf.1	N	1 - o1- 165;tml 66- 188;i18 9-194;		ASN_GLYCOSYLATION 128-131; AMIDATION 190- 193; MYRISTYL 170-175; CK2_PHOSPHO_SITE 3-6; PKC_PHOSPHO_SITE 67-69; AMIDATION 3-6; ASN_GLYCOSYLATION 92-95; MYRISTYL 157-162; CK2_PHOSPHO_SITE 130-133; MYRISTYL 184-189; PKC_PHOSPHO_SITE 3-5; PKC_PHOSPHO_SITE 98-100; ASN_GLYCOSYLATION 20-23; MYRISTYL 7-12; MYRISTYL 180-185; PKC_PHOSPHO_SITE 160-162; CK2_PHOSPHO_SITE 7-10; ASN_GLYCOSYLATION 6-9; CAMP_PHOSPHO_SITE 5-8; MYRISTYL 2- 7;	MHCCLASSI 64-80; MHCCLASSI 135-153;
DEX0449_03 1.aa.1	N	0 - o1- 132;	122- 129,1.15; 10-24,1.133; 67-74,1.074; 106-	MYRISTYL 22-27; PKC_PHOSPHO_SITE 33-35; ASN_GLYCOSYLATION 17-20;	

			117,1.25; 94- 102,1.038;	ASN_GLYCOSYLATION 2-5; PKC_PHOSPHO_SITE 118-120; MYRISTYL 61-66; PKC_PHOSPHO_SITE 4-6; CK2_PHOSPHO_SITE 4-7; AMIDATION 103-106; MYRISTYL 89-94; CAMP_PHOSPHO_SITE 76-79; PKC_PHOSPHO_SITE 32-34; CK2_PHOSPHO_SITE 88-91; ASN_GLYCOSYLATION 27-30;	
DEX0449_03 1.orf.1	N	0 - 01- 199;	141- 146,1.033; 46-52,1.027; 54-61,1.105; 153- 163,1.116; 171- 192,1.152; 94- 101,1.084;	CK2_PHOSPHO_SITE 192-195; PKC_PHOSPHO_SITE 166-168; PKC_PHOSPHO_SITE 3-5; PKC_PHOSPHO_SITE 148-150; PKC_PHOSPHO_SITE 102-104; PKC_PHOSPHO_SITE 20-22; PKC_PHOSPHO_SITE 184-186; PKC_PHOSPHO_SITE 50-52; CK2_PHOSPHO_SITE 148-151; ASN_GLYCOSYLATION 167-170; CK2_PHOSPHO_SITE 79-82; CAMP_PHOSPHO_SITE 12-15; ASN_GLYCOSYLATION 77-80; PKC_PHOSPHO_SITE 79-81;	LYS_RICH 5-40;
DEX0449_03 2.aa.1	N	0 - 01- 259;	122- 131,1.066; 136- 174,1.19; 100- 106,1.068; 82-89,1.039; 190- 197,1.171; 232- 238,1.086; 204-	ASN_GLYCOSYLATION 130-133; MYRISTYL 98-103; MYRISTYL 53-58; CK2_PHOSPHO_SITE 132-135; MYRISTYL 52-57; AMIDATION 2-5; PKC_PHOSPHO_SITE 250-252; PKC_PHOSPHO_SITE 45-47; MYRISTYL	UBCc 111-255; UQ_con 107-250; UBIQUITIN_CONJUGAT_ 1 181-195; sp_O14933_UBC8_HUMA N 111-250; UBIQUITIN_CONJUGAT_ 2 111-244;

			220,1.175; 177- 184,1.082; 10-16,1.056; 21-53,1.139;	17-22; MYRISTYL 65-70; CK2_PHOSPHO_SITE 115-118; PKC_PHOSPHO_SITE 177-179; MYRISTYL 74-79; CK2_PHOSPHO_SITE 93-96; PKC_PHOSPHO_SITE 93-95; CK2_PHOSPHO_SITE 113-116; MYRISTYL 70-75; MYRISTYL 69-74;	
DEX0449_03 2.aa.2	N	0 - i1- 156;	87-94,1.171; 129- 135,1.086; 74-81,1.082; 101- 117,1.175; 19-28,1.066; 33-71,1.19;	PKC_PHOSPHO_SITE 147-149; PKC_PHOSPHO_SITE 74-76; CAMP_PHOSPHO_SITE 2-5; ASN_GLYCOSYLATION 27-30; CK2_PHOSPHO_SITE 29-32; PKC_PHOSPHO_SITE 5-7; CK2_PHOSPHO_SITE 12-15;	UBCc 8-152; UBIQUITIN_CONJUGAT_ 1 78-92; UBIQUITIN_CONJUGAT_ 2 13-141; sp_O14933_UBC8_HUMA N 13-147; UQ_con 11-147;
DEX0449_03 2.orf.2	N	0 - o1- 168;	31-40,1.066; 45-83,1.19; 4-10,1.1; 99- 106,1.171; 141- 147,1.086; 86-93,1.082; 113- 129,1.175;	PKC_PHOSPHO_SITE 159-161; CK2_PHOSPHO_SITE 41-44; ASN_GLYCOSYLATION 39-42; MYRISTYL 16-21; PKC_PHOSPHO_SITE 86-88;	UBIQUITIN_CONJUGAT_ 2 20-153; UBCc 20- 164; UQ_con 15-159; sp_O14933_UBC8_HUMA N 25-159; UBIQUITIN_CONJUGAT_ 1 90-104;
DEX0449_03 2.aa.3	N	0 - o1- 137;	82-98,1.175; 68-75,1.171; 110- 116,1.086; 14-52,1.19; 55-62,1.082;	AMIDATION 2-5; PKC_PHOSPHO_SITE 55-57; ASN_GLYCOSYLATION 8-11; PKC_PHOSPHO_SITE 128-130; MYRISTYL 7-12; CK2_PHOSPHO_SITE 10-13;	UBCc 3-133; UBIQUITIN_CONJUGAT_ 2 12-122; UBIQUITIN_CONJUGAT_ 1 59-73; sp_O14933_UBC8_HUMA N 8-128; UQ_con 1- 128;
DEX0449_03 2.aa.4	N	0 - o1- 237;	155- 162,1.082; 113- 118,1.041; 168- 175,1.171; 121- 152,1.095; 75-82,1.124;	ASN_GLYCOSYLATION 24-27; PKC_PHOSPHO_SITE 65-67; MYRISTYL 85-90; PKC_PHOSPHO_SITE 155-157; PKC_PHOSPHO_SITE 4-6;	sp_O14933_UBC8_HUMA N 126-228; sp_O14933_UBC8_HUMA N 3-42; UQ_con 1- 228; UBIQUITIN_CONJUGAT_ 1 159-173; UBIQUITIN_CONJUGAT_ 2 122-222; UBCc 5-

			95-101,1.1; 30-58,1.19; 182- 198,1.175; 210- 216,1.086; 16-25,1.066; 5-14,1.074;	PKC_PHOSPHO_SITE 228-230; CK2_PHOSPHO_SITE 26-29; PKC_PHOSPHO_SITE 118-120;	233;
DEX0449_03 2.orf.4	N	0 - o1- 156;	32-37,1.041; 129- 135,1.086; 101- 117,1.175; 74-81,1.082; 87-94,1.171; 40-71,1.095;	CK2_PHOSPHO_SITE 16-19; PKC_PHOSPHO_SITE 147-149; PKC_PHOSPHO_SITE 74-76; PKC_PHOSPHO_SITE 37-39; MYRISTYL 14-19;	UBIQUITIN_CONJUGAT_ 2 41-141; UQ_con 14-147; UBIQUITIN_CONJUGAT_ 1 78-92; UBCC 12- 152; sp_O14933_UBC8_HUMA N 45-147;
DEX0449_03 3.aa.1	N	0 - o1- 88;	51-58,1.097; 4-12,1.097; 23-30,1.121;	PKC_PHOSPHO_SITE 38-40; CK2_PHOSPHO_SITE 48-51; ASN_GLYCOSYLATION 79-82;	
DEX0449_03 3.aa.2	N	0 - o1- 40;	21-27,1.087; 31-37,1.071;	MYRISTYL 7-12; AMIDATION 10-13;	
DEX0449_03 3.orf.2	N	0 - o1- 60;	21-39,1.171; 11-17,1.065;	MYRISTYL 17-22; MYRISTYL 55-60;	
DEX0449_03 4.aa.1	N	0 - o1- 213;		PKC_PHOSPHO_SITE 77-79; CK2_PHOSPHO_SITE 64-67; MYRISTYL 108-113; PKC_PHOSPHO_SITE 209-211; CK2_PHOSPHO_SITE 161-164; CK2_PHOSPHO_SITE 122-125; MYRISTYL 23-28; CK2_PHOSPHO_SITE 113-116; MYRISTYL 5-10; CK2_PHOSPHO_SITE 77-80; PKC_PHOSPHO_SITE 64-66; MYRISTYL 73-78; CK2_PHOSPHO_SITE 209-212;	
DEX0449_03 5.aa.1	N	0 - o1- 458;	223- 245,1.186; 187- 196,1.158; 175- 180,1.057; 427- 436,1.181; 279-	ASN_GLYCOSYLATION 328-331; MYRISTYL 243-248; PKC_PHOSPHO_SITE 352-354; ASN_GLYCOSYLATION 171-174; PKC_PHOSPHO_SITE 420-422;	ANK 352-381; ANK 219-248; ANK 388- 421; ank 291-350; ANK_REPEAT_3 388- 424; ank 183-215; ank 388-424; ANK 183-212; ANK 291- 321; ANKYRIN 405- 417; ANK REPEAT 2

			286,1.076; 256- 263,1.17; 28-61,1.087; 72-85,1.167; 366- 385,1.175; 198- 205,1.169; 108- 114,1.091; 293- 310,1.204; 316- 347,1.133; 119- 125,1.087; 154- 162,1.094; 265- 276,1.128; 355- 361,1.101; 392- 412,1.161; 100- 106,1.066;	CK2_PHOSPHO_SITE 330-333; ASN_GLYCOSYLATION 149-152; ASN_GLYCOSYLATION 77-80; AMIDATION 194-197; ASN_GLYCOSYLATION 247-250; PKC_PHOSPHO_SITE 43-45; MYRISTYL 275-280; MYRISTYL 416-421; CK2_PHOSPHO_SITE 120-123; MYRISTYL 145-150; ASN_GLYCOSYLATION 136-139; CK2_PHOSPHO_SITE 288-291; CK2_PHOSPHO_SITE 18-21; PKC_PHOSPHO_SITE 334-336; CK2_PHOSPHO_SITE 404-407;	352-384; ank 352- 384; ANK_REPEAT_1 219-251; ANK_REP_REGION 183- 432; ank 252-290; ank 219-251; ANKYRIN 253-265; ANK 252-283;
DEX0449_03 5.orf.1	N	0 - 01- 439;	274- 291,1.204; 100- 106,1.087; 168- 177,1.158; 179- 186,1.169; 297- 328,1.133; 4-42,1.154; 246- 257,1.128; 204- 226,1.186; 336- 342,1.101; 347- 366,1.175; 53-66,1.167; 260- 267,1.076; 135- 143,1.094; 156- 161,1.057; 237- 244,1.17; 408- 417,1.181; 81-87,1.066; 373-	ASN_GLYCOSYLATION 309-312; ASN_GLYCOSYLATION 130-133; MYRISTYL 224-229; ASN_GLYCOSYLATION 152-155; ASN_GLYCOSYLATION 228-231; MYRISTYL 256-261; CK2_PHOSPHO_SITE 101-104; AMIDATION 175- 178; PKC_PHOSPHO_SITE 315-317; PKC_PHOSPHO_SITE 333-335; MYRISTYL 126-131; MYRISTYL 397-402; PKC_PHOSPHO_SITE 401-403; ASN_GLYCOSYLATION 58-61; CK2_PHOSPHO_SITE 311-314; PKC_PHOSPHO_SITE 24-26; CK2_PHOSPHO_SITE 385-388; CK2_PHOSPHO_SITE 269-272;	ANK 272-302; ANK 369-402; ANK_REPEAT_1 200- 232; ANKYRIN 234- 246; ANKYRIN 386- 398; ANK_REPEAT_3 369-405; ANK 233- 264; ANK 200-229; ank 272-331; ANK 164-193; ANK_REP_REGION 164- 413; ank 200-232; ANK_REPEAT_2 333- 365; ank 333-365; ank 164-196; ank 369-405; ank 233- 271; ANK 333-362;

			393,1.161; 89-95,1.091;	ASN_GLYCOSYLATION 117-120;	
DEX0449_03 6.aa.1	Y	0 - 01- 130;	4-21,1.165; 23-43,1.129; 110- 127,1.174; 49-67,1.178; 82- 102,1.143;	CK2_PHOSPHO_SITE 68-71; PKC_PHOSPHO_SITE 81-83; AMIDATION 74-77; MYRISTYL 19-24; CK2_PHOSPHO_SITE 81-84; CAMP_PHOSPHO_SITE 76-79; PKC_PHOSPHO_SITE 127-129;	
DEX0449_03 6.orf.1	N	0 - 01- 243;	157- 196,1.173; 83-93,1.055; 215- 223,1.167; 11-37,1.163; 139- 149,1.056; 225- 240,1.175; 110- 116,1.067; 49-61,1.064; 119- 131,1.061; 201- 208,1.12;	MYRISTYL 135-140; PKC_PHOSPHO_SITE 173-175; MYRISTYL 19-24; MYRISTYL 61-66; MYRISTYL 121-126; MYRISTYL 125-130; MYRISTYL 46-51; CK2_PHOSPHO_SITE 196-199; MYRISTYL 151-156; MYRISTYL 37-42; PKC_PHOSPHO_SITE 206-208; MYRISTYL 33-38; MYRISTYL 50-55; CK2_PHOSPHO_SITE 206-209;	TGF_BETA_2 145-231; sp_Q64280_TGF4_MOUSE E 145-230; TGF-beta 137-231; TGF_BETA_1 159-174; TGFB 141- 231;
DEX0449_03 7.aa.1	N	0 - 01- 204;	152- 164,1.099; 138- 144,1.057; 194- 201,1.064; 4-21,1.091; 30-61,1.19; 125- 131,1.074; 63- 113,1.187;	CK2_PHOSPHO_SITE 113-116; CK2_PHOSPHO_SITE 28-31; MYRISTYL 153-158; MYRISTYL 134-139; CK2_PHOSPHO_SITE 109-112; ASN_GLYCOSYLATION 59-62; MYRISTYL 9-14; PKC_PHOSPHO_SITE 163-165; CK2_PHOSPHO_SITE 141-144;	
DEX0449_03 8.aa.1	Y	0 - 01- 80;	4-36,1.28; 67-77,1.06;	CK2_PHOSPHO_SITE 64-67; MYRISTYL 71-76;	
DEX0449_03 8.orf.2	N	0 - 01- 84;	15-21,1.024; 65-71,1.084; 29-37,1.061;	CK2_PHOSPHO_SITE 59-62; PKC_PHOSPHO_SITE 12-14; PKC_PHOSPHO_SITE 41-43; CK2_PHOSPHO_SITE 60-63;	

				CK2_PHOSPHO_SITE 5-8; CAMP_PHOSPHO_SITE 74-77; PKC_PHOSPHO_SITE 22-24;	
DEX0449_03 8.aa.3	Y	0 - 01- 56;	4-17,1.159; 34-53,1.146;		
DEX0449_03 9.aa.1	Y	0 - 01- 451;	26-42,1.143; 70-77,1.072; 4-11,1.207; 410- 415,1.049; 181- 193,1.095; 88- 100,1.087; 368- 381,1.12; 328- 341,1.121; 62-68,1.081; 200- 260,1.115; 16-24,1.105; 46-60,1.115; 305- 311,1.064; 264- 277,1.08; 433- 439,1.075; 112- 118,1.13;	MYRISTYL 351-356; MYRISTYL 365-370; CK2_PHOSPHO_SITE 82-85; CAMP_PHOSPHO_SITE 360-363; PKC_PHOSPHO_SITE 389-391; CK2_PHOSPHO_SITE 336-339; MYRISTYL 415-420; MYRISTYL 106-111; CK2_PHOSPHO_SITE 160-163; PKC_PHOSPHO_SITE 322-324; CK2_PHOSPHO_SITE 125-128; CK2_PHOSPHO_SITE 20-23; MYRISTYL 254-259; CK2 PHOSPHO_SITE 414-417; CK2_PHOSPHO_SITE 294-297; CK2_PHOSPHO_SITE 96-99; LEUCINE_ZIPPER 232-253; ASN_GLYCOSYLATION 158-161; TYR_PHOSPHO_SITE 244-251; MYRISTYL 348-353; CK2_PHOSPHO_SITE 166-169; CK2_PHOSPHO_SITE 129-132; MYRISTYL 438-443; CK2_PHOSPHO_SITE 143-146; AMIDATION 75-78; MYRISTYL 407-412; MYRISTYL 349-354; ASN_GLYCOSYLATION 164-167; CK2_PHOSPHO_SITE 261-264; PKC_PHOSPHO_SITE 359-361;	TPR 45-78; TPR 45- 78; TPR 247-280; TPR_REPEAT_1 45-78; TPR_REPEAT_2 205- 238; TPR 247-280; TPR 205-238; GLU_RICH 113-321; ATP_GTP_A 89-96; TPR 205-238; TPR_REGION 205-280; TPR_REPEAT_3 247- 280;
DEX0449 04	N	0 - 01-	189-	CK2 PHOSPHO SITE	gmd 24-331;

0.aa.1		337;	208,1.136; 4-11,1.092; 65-75,1.046; 291- 301,1.06; 313- 319,1.127; 251- 260,1.089; 235-249,1.1; 159- 165,1.066; 89-100,1.17; 24-30,1.112; 218- 224,1.061; 279- 289,1.183; 172- 178,1.061; 36-57,1.131; 107- 157,1.134;	292-295; MYRISTYL 198-203; CK2_PHOSPHO_SITE 32-35; PKC_PHOSPHO_SITE 277-279; MYRISTYL 143-148; PKC_PHOSPHO_SITE 62-64; TYR_PHOSPHO_SITE 115-123; PKC_PHOSPHO_SITE 10-12; CAMP_PHOSPHO_SITE 186-189; CAMP_PHOSPHO_SITE 55-58; MYRISTYL 13-18; PKC_PHOSPHO_SITE 132-134; CK2_PHOSPHO_SITE 154-157; PKC_PHOSPHO_SITE 245-247; MYRISTYL 206-211; MYRISTYL 33-38; CK2_PHOSPHO_SITE 245-248; PKC_PHOSPHO_SITE 185-187; MYRISTYL 36-41; PKC_PHOSPHO_SITE 147-149;	
DEX0449_04 0.aa.2	N	0 - 01- 332;	24-30,1.112; 314- 324,1.183; 36-57,1.131; 172- 194,1.171; 253- 259,1.061; 107- 157,1.134; 4-11,1.092; 286- 295,1.089; 159- 165,1.066; 224- 243,1.136; 65-75,1.046; 89-100,1.17; 270-284,1.1; 196- 209,1.113;	CK2_PHOSPHO_SITE 32-35; MYRISTYL 204-209; CAMP_PHOSPHO_SITE 55-58; PKC_PHOSPHO_SITE 147-149; MYRISTYL 241-246; TYR_PHOSPHO_SITE 115-123; PKC_PHOSPHO_SITE 62-64; CK2_PHOSPHO_SITE 280-283; PKC_PHOSPHO_SITE 10-12; PKC_PHOSPHO_SITE 132-134; MYRISTYL 143-148; CK2_PHOSPHO_SITE 154-157; MYRISTYL 13-18; PKC_PHOSPHO_SITE 211-213; MYRISTYL 233-238; PKC PHOSPHO SITE	ADH_SHORT 166-194; gmd 24-332;

				312-314; CAMP_PHOSPHO_SITE 221-224; MYRISTYL 33-38; PKC_PHOSPHO_SITE 280-282; CK2_PHOSPHO_SITE 327-330; MYRISTYL 36-41; PKC_PHOSPHO_SITE 220-222;	
DEX0449_04 0.aa.3	N	0 - ol- 382;	172- 194,1.171; 196- 209,1.113; 286- 295,1.089; 159- 165,1.066; 253- 259,1.061; 314- 324,1.183; 343- 352,1.075; 224- 243,1.136; 107- 157,1.134; 36-57,1.131; 65-75,1.046; 4-11,1.092; 24-30,1.112; 357- 374,1.131; 326- 341,1.108; 89-100,1.17; 270-284,1.1;	PKC_PHOSPHO_SITE 147-149; PKC_PHOSPHO_SITE 132-134; CK2_PHOSPHO_SITE 32-35; MYRISTYL 33-38; PKC_PHOSPHO_SITE 10-12; MYRISTYL 204-209; PKC_PHOSPHO_SITE 280-282; CK2_PHOSPHO_SITE 154-157; PKC_PHOSPHO_SITE 62-64; MYRISTYL 352-357; MYRISTYL 13-18; CK2_PHOSPHO_SITE 280-283; PKC_PHOSPHO_SITE 220-222; TYR_PHOSPHO_SITE 115-123; MYRISTYL 143-148; CAMP_PHOSPHO_SITE 55-58; MYRISTYL 241-246; PKC_PHOSPHO_SITE 211-213; PKC_PHOSPHO_SITE 312-314; CAMP_PHOSPHO_SITE 221-224; MYRISTYL 233-238; MYRISTYL 36-41;	gmd 24-361; ADH_SHORT 166-194;
DEX0449_04 1.aa.1	N	0 - ol- 258;	41-48,1.083; 245- 255,1.126; 147- 154,1.063; 134- 139,1.059; 108- 124,1.184; 92-99,1.184; 6-23,1.154; 212-	PKC_PHOSPHO_SITE 232-234; MYRISTYL 159-164; AMIDATION 232- 235; CAMP_PHOSPHO_SITE 241-244; CK2_PHOSPHO_SITE 178-181; MYRISTYL 174-179; PKC_PHOSPHO_SITE 115-117;	UQ_con 9-154; sp_Q9BTC1_Q9BTC1_HU MAN 28-153; UBCC 14-157; UBIQUITIN_CONJUGAT_ 1 84-98; UBIQUITIN_CONJUGAT_ 2 14-146;

			220,1.139; 61-69,1.073; 168- 174,1.029; 189- 197,1.112; 71-88,1.083;	CK2_PHOSPHO_SITE 173-176; MYRISTYL 52-57; PKC_PHOSPHO_SITE 162-164; AMIDATION 185- 188; CAMP_PHOSPHO_SITE 205-208; MYRISTYL 247-252; PKC_PHOSPHO_SITE 240-242; CK2_PHOSPHO_SITE 3-6; AMIDATION 203-206; PKC_PHOSPHO_SITE 244-246; ASN_GLYCOSYLATION 225-228;	
DEX0449_04 1.aa.2	N	0 - ol- 262;	226- 232,1.029; 166- 182,1.184; 23-30,1.141; 205- 212,1.063; 192- 197,1.059; 99- 106,1.083; 65-81,1.154; 129- 146,1.083; 47-60,1.082; 119- 127,1.073; 4-18,1.094; 150- 157,1.184;	MYRISTYL 61-66; CK2_PHOSPHO_SITE 231-234; MYRISTYL 248-253; AMIDATION 31-34; MYRISTYL 232-237; PKC_PHOSPHO_SITE 173-175; MYRISTYL 110-115; MYRISTYL 247-252; AMIDATION 41-44; CK2_PHOSPHO_SITE 236-239; MYRISTYL 217-222; MYRISTYL 45-50; MYRISTYL 24-29; MYRISTYL 38-43; PKC_PHOSPHO_SITE 220-222; MYRISTYL 244-249; MYRISTYL 254-259;	sp_Q9BTC1_Q9BTC1_HU MAN 86-211; UBCC 72-215; UBIQUITIN_CONJUGAT_ 2 72-204; UBIQUITIN_CONJUGAT_ 1 142-156; UQ_con 67-212;
DEX0449_04 1.orf.2	N	0 - ol- 291;	140- 157,1.083; 34-41,1.141; 177- 193,1.184; 58-71,1.082; 76-92,1.154; 110- 117,1.083; 216- 223,1.063; 161- 168,1.184; 12-29,1.094; 130- 138,1.073; 237- 243,1.029; 203-	MYRISTYL 255-260; MYRISTYL 56-61; CK2_PHOSPHO_SITE 247-250; PKC_PHOSPHO_SITE 231-233; PKC_PHOSPHO_SITE 282-284; MYRISTYL 49-54; CK2_PHOSPHO_SITE 242-245; MYRISTYL 72-77; PKC_PHOSPHO_SITE 184-186; AMIDATION 52-55; MYRISTYL 121-126; AMIDATION 42-45; MYRISTYL 35-40; MYRISTYL 243-248;	UBCC 83-226; UQ_con 78-223; sp_Q9BTC1_Q9BTC1_HU MAN 97-222; ARG RICH 5-59; UBIQUITIN_CONJUGAT_ 1 153-167; UBIQUITIN_CONJUGAT_ 2 83-215;

			208,1.059;	MYRISTYL 228-233; CAMP_PHOSPHO_SITE 279-282;	
DEX0449_04 1.aa.3	N	0 - 01- 233;	163- 168,1.059; 6-23,1.154; 41-48,1.083; 197- 203,1.029; 92-99,1.184; 140- 153,1.184; 176- 183,1.063; 108- 117,1.18; 61-69,1.073; 71-88,1.083; 119- 137,1.121;	MYRISTYL 218-223; PKC_PHOSPHO_SITE 116-118; CK2_PHOSPHO_SITE 207-210; MYRISTYL 215-220; CK2_PHOSPHO_SITE 202-205; PKC_PHOSPHO_SITE 144-146; CK2_PHOSPHO_SITE 3-6; MYRISTYL 52- 57; MYRISTYL 219- 224; CK2_PHOSPHO_SITE 133-136; PKC_PHOSPHO_SITE 191-193; MYRISTYL 188-193; CK2_PHOSPHO_SITE 116-119; MYRISTYL 203-208; MYRISTYL 225-230;	UBIQUITIN_CONJUGAT_ 2 14-175; UBIQUITIN_CONJUGAT_ 1 84-98; sp_Q9BTC1_Q9BTC1_HU MAN 28-182; UBCc 14-186; UQ_con 9- 183;
DEX0449_04 1.orf.3	N	0 - 01- 390;	145- 162,1.154; 279- 292,1.184; 336- 342,1.029; 180- 187,1.083; 315- 322,1.063; 231- 238,1.184; 200- 208,1.073; 247- 256,1.18; 15-31,1.195; 41-51,1.11; 65-90,1.129; 210- 227,1.083; 258- 276,1.121; 302- 307,1.059;	MYRISTYL 354-359; AMIDATION 80-83; PKC_PHOSPHO_SITE 283-285; CK2_PHOSPHO_SITE 346-349; CK2_PHOSPHO_SITE 255-258; PKC_PHOSPHO_SITE 255-257; MYRISTYL 128-133; CK2_PHOSPHO_SITE 142-145; MYRISTYL 342-347; AMIDATION 118- 121; CAMP_PHOSPHO_SITE 90-93; MYRISTYL 191-196; CAMP_PHOSPHO_SITE 378-381; PKC_PHOSPHO_SITE 381-383; PKC_PHOSPHO_SITE 33-35; MYRISTYL 85-90; CK2_PHOSPHO_SITE 272-275; PKC_PHOSPHO_SITE 45-47; PKC_PHOSPHO_SITE 330-332; CK2 PHOSPHO SITE	OXYTOCINR 324-337; UBCc 153-325; UBIQUITIN_CONJUGAT_ 2 153-314; UBIQUITIN_CONJUGAT_ 1 223-237; UQ_con 148-322; sp_Q9BTC1_Q9BTC1_HU MAN 167-321; OXYTOCINR 93-106;

				341-344; MYRISTYL 327-332;	
DEX0449_04 2.aa.1	N	0 - ol- 110;	59-66,1.064; 23-35,1.2; 87- 107,1.135; 43-51,1.055; 4-19,1.144; 71-80,1.11;	PKC_PHOSPHO_SITE 92-94; CAMP_PHOSPHO_SITE 37-40; MYRISTYL 2-7; PKC_PHOSPHO_SITE 35-37; MYRISTYL 56-61;	
DEX0449_04 3.aa.1	N	1 - il- 95;tm96 - 118;ol1 9-257;	48-55,1.06; 228- 234,1.108; 249- 254,1.068; 202- 210,1.12; 129- 146,1.215; 62-72,1.097; 212- 224,1.14; 157- 169,1.121; 185- 191,1.092; 94- 118,1.175; 76-82,1.04; 171- 180,1.089; 240- 247,1.039; 22-34,1.191;	MYRISTYL 9-14; PKC_PHOSPHO_SITE 214-216; PKC_PHOSPHO_SITE 148-150; CK2_PHOSPHO_SITE 210-213; PKC_PHOSPHO_SITE 184-186; CK2_PHOSPHO_SITE 201-204; PKC_PHOSPHO_SITE 224-226;	AhpC-TSA 67-216;
DEX0449_04 4.aa.1	N	0 - il- 105;	74-81,1.122; 95- 102,1.095; 24-48,1.119;	CAMP_PHOSPHO_SITE 4-7; AMIDATION 16-19; CAMP_PHOSPHO_SITE 63-66; PKC_PHOSPHO_SITE 91-93; ASN_GLYCOSYLATION 21-24; ASN_GLYCOSYLATION 56-59; PKC_PHOSPHO_SITE 69-71; ASN_GLYCOSYLATION 65-68; PKC_PHOSPHO_SITE 58-60;	sp_P02403_RL37_HUMA N 12-62; Ribosomal_L37e 10- 63; RIBOSOMAL_L37E 12-31;
DEX0449_04 4.orf.1	N	0 - ol- 111;	30-54,1.119; 80-87,1.122; 102- 108,1.095;	PKC_PHOSPHO_SITE 75-77; ASN_GLYCOSYLATION 71-74; ASN_GLYCOSYLATION 27-30; AMIDATION 22-25;	Ribosomal_L37e 18- 69; sp_P02403_RL37_HUMA N 18-68; RIBOSOMAL_L37E 18- 37;

				ASN_GLYCOSYLATION 62-65; PKC_PHOSPHO_SITE 64-66; CAMP_PHOSPHO_SITE 69-72; PKC_PHOSPHO_SITE 97-99;	
DEX0449_04 4.aa.2	N	0 - 01- 104;		ASN_GLYCOSYLATION 20-23; PKC_PHOSPHO_SITE 2-4; ASN_GLYCOSYLATION 55-58; PKC_PHOSPHO_SITE 57-59; AMIDATION 15-18; CK2_PHOSPHO_SITE 2-5; CAMP_PHOSPHO_SITE 62-65; PKC_PHOSPHO_SITE 4-6; PKC_PHOSPHO_SITE 90-92; PKC_PHOSPHO_SITE 68-70; ASN_GLYCOSYLATION 64-67;	Ribosomal_L37e 9- 62; sp_P02403_RL37_HUMA N 10-61; RIBOSOMAL_L37E 11- 30;
DEX0449_04 4.orf.2	N	0 - 01- 129;		MYRISTYL 31-36; PKC_PHOSPHO_SITE 94-96; PKC_PHOSPHO_SITE 2-4; MYRISTYL 50- 55; ASN_GLYCOSYLATION 35-38; PKC_PHOSPHO_SITE 108-110; ASN_GLYCOSYLATION 12-15; PKC_PHOSPHO_SITE 11-13; CK2_PHOSPHO_SITE 62-65; CK2_PHOSPHO_SITE 108-111; PKC_PHOSPHO_SITE 20-22; PKC_PHOSPHO_SITE 10-12;	
DEX0449_04 4.aa.3	N	0 - i1- 96;	15-39,1.104; 87-93,1.095; 4-9,1.129; 65-72,1.122;	ASN_GLYCOSYLATION 56-59; CAMP_PHOSPHO_SITE 54-57; ASN_GLYCOSYLATION 47-50; PKC_PHOSPHO_SITE 82-84;	Ribosomal_L37e 7- 54; sp_P02403_RL37_HUMA N 12-53;

				PKC_PHOSPHO_SITE 60-62; PKC_PHOSPHO_SITE 49-51;	
DEX0449_04 4.aa.4	N	0 - 11- 55;	23-36,1.118; 39-45,1.072;	PKC_PHOSPHO_SITE 20-22; CK2_PHOSPHO_SITE 34-37; PKC_PHOSPHO_SITE 34-36; CK2_PHOSPHO_SITE 7-10;	
DEX0449_04 4.orf.4	N	0 - 01- 72;	63-69,1.095; 4-15,1.104; 41-48,1.122;	PKC_PHOSPHO_SITE 58-60; CAMP_PHOSPHO_SITE 30-33; PKC_PHOSPHO_SITE 25-27; PKC_PHOSPHO_SITE 36-38; ASN_GLYCOSYLATION 23-26; ASN_GLYCOSYLATION 32-35;	Ribosomal_L37e 1- 30; sp_Q9D823_Q9D823_MO USE 1-29;
DEX0449_04 5.aa.1	N	0 - 01- 185;	7-13,1.077; 21-27,1.045; 65-72,1.166; 42-56,1.188; 137- 162,1.156; 169- 175,1.06; 100- 106,1.164; 110- 127,1.193;	ASN_GLYCOSYLATION 98-101; MYRISTYL 9-14; MYRISTYL 178-183; ASN_GLYCOSYLATION 61-64; CK2_PHOSPHO_SITE 106-109; PKC_PHOSPHO_SITE 151-153; MYRISTYL 14-19; PKC_PHOSPHO_SITE 127-129; PKC_PHOSPHO_SITE 77-79; MYRISTYL 19-24;	
DEX0449_04 5.orf.1	N	0 - 01- 214;	194- 211,1.128; 14-21,1.059; 182- 191,1.098; 86- 101,1.148; 49-67,1.235; 72-83,1.199; 105- 113,1.105; 120- 180,1.17; 29-47,1.107; 4-11,1.148;	MYRISTYL 71-76; ASN_GLYCOSYLATION 147-150; MYRISTYL 145-150; PKC_PHOSPHO_SITE 176-178; MYRISTYL 150-155; PKC_PHOSPHO_SITE 66-68; PKC_PHOSPHO_SITE 34-36;	
DEX0449_04 5.aa.2	N	0 - 01- 131;	7-13,1.077; 21-27,1.045; 100-	ASN_GLYCOSYLATION 61-64; ASN_GLYCOSYLATION	RIBOSOMAL_L18E 70- 87; sp_O07020_RL18 HUMA

			106,1.164; 65-72,1.166; 42-56,1.188; 110- 122,1.136;	98-101; MYRISTYL 19-24; CK2_PHOSPHO_SITE 106-109; PKC_PHOSPHO_SITE 77-79; MYRISTYL 9-14; MYRISTYL 14-19;	N 49-119; Ribosomal_L18e 31- 131;
DEX0449_04 5.orf.2	N	0 - o1- 132;	22-28,1.045; 101- 107,1.164; 66-73,1.166; 9-14,1.005; 43-57,1.188; 111- 123,1.136;	ASN_GLYCOSYLATION 62-65; ASN_GLYCOSYLATION 99-102; MYRISTYL 15-20; CK2_PHOSPHO_SITE 107-110; MYRISTYL 20-25; MYRISTYL 10-15; MYRISTYL 6-11; PKC_PHOSPHO_SITE 78-80;	RIBOSOMAL_L18E 71- 88; Ribosomal_L18e 32-132; sp_Q07020_RL18_HUMA N 50-120;
DEX0449_04 5.aa.3	N	0 - o1- 195;	180- 191,1.088; 143- 153,1.193; 74-80,1.087; 14-20,1.045; 58-65,1.166; 94- 100,1.115; 35-49,1.188; 160- 172,1.133; 115- 141,1.177;	ASN_GLYCOSYLATION 54-57; MYRISTYL 7-12; MYRISTYL 172-177; MYRISTYL 127-132; MYRISTYL 12-17; MYRISTYL 109-114; MYRISTYL 189-194; PKC_PHOSPHO_SITE 4-6; PKC_PHOSPHO_SITE 70-72;	sp_Q9D987_Q9D987_MO USE 42-81; Ribosomal_L18e 24- 160; RIBOSOMAL_L18E 63-80;
DEX0449_04 6.aa.1	Y	0 - o1- 115;	18-51,1.229; 54- 109,1.225;	PKC_PHOSPHO_SITE 107-109; MYRISTYL 34-39; MYRISTYL 102-107; PKC_PHOSPHO_SITE 38-40; MYRISTYL 98-103; TYR_PHOSPHO_SITE 40-48;	
DEX0449_04 6.aa.2	N	0 - o1- 156;	91- 126,1.217; 37-83,1.178; 133- 149,1.145;	MYRISTYL 126-131; MYRISTYL 6-11; CK2_PHOSPHO_SITE 30-33; ASN_GLYCOSYLATION 28-31; MYRISTYL 95-100; MYRISTYL 38-43; PKC_PHOSPHO_SITE 150-152; PKC_PHOSPHO_SITE 9-11; MYRISTYL 17-22; MYRISTYL 14-19; PKC PHOSPHO SITE	TIMP 13-148; NTR 26-148;

				127-129;	
DEX0449_04 7.aa.1	Y	0 - 01- 226;	141- 154,1.109; 159- 176,1.17; 180- 210,1.214; 42-57,1.155; 212- 223,1.121; 76- 133,1.186; 4-22,1.064; 32-40,1.093;	MYRISTYL 26-31; CK2_PHOSPHO_SITE 71-74; MYRISTYL 21-26; MYRISTYL 9-14; CK2_PHOSPHO_SITE 132-135; CK2_PHOSPHO_SITE 73-76; MYRISTYL 27-32; CK2_PHOSPHO_SITE 67-70; MYRISTYL 34-39; MYRISTYL 44-49;	CYCLIN 83-185; KV14CHANNEL 19-28; ALA_RICH 2-22; ANTIFREEZEI 13-22; ANTIFREEZEI 1-11; KV14CHANNEL 6-16; cyclin 48-192;
DEX0449_04 7.orf.1	N	0 - 01- 253;	103- 125,1.18; 141- 154,1.056; 4-13,1.152; 176- 182,1.102; 158- 167,1.11; 37-48,1.194; 57-88,1.216; 90-96,1.106; 15-24,1.127; 197- 203,1.06;	PKC_PHOSPHO_SITE 229-231; CK2_PHOSPHO_SITE 167-170; CAMP_PHOSPHO_SITE 232-235; PKC_PHOSPHO_SITE 184-186; ASN_GLYCOSYLATION 53-56; MYRISTYL 143-148; MYRISTYL 30-35; PKC_PHOSPHO_SITE 180-182; TYR_PHOSPHO_SITE 214-220; PKC_PHOSPHO_SITE 11-13; MYRISTYL 202-207; PKC_PHOSPHO_SITE 246-248; CK2_PHOSPHO_SITE 98-101; MYRISTYL 238-243; PKC_PHOSPHO_SITE 167-169; MYRISTYL 242-247; CK2_PHOSPHO_SITE 121-124; PKC_PHOSPHO_SITE 55-57; MYRISTYL 190-195;	SER_RICH 209-249; CYCLIN 13-111;
DEX0449_04 7.orf.2	N	0 - 01- 346;	57-88,1.216; 176- 182,1.102; 103- 125,1.18; 15-24,1.127; 197- 203,1.06; 4- 13,1.152; 37-48,1.194; 280- 287.1.082;	PKC_PHOSPHO_SITE 180-182; TYR_PHOSPHO_SITE 276-283; PKC_PHOSPHO_SITE 333-335; MYRISTYL 190-195; PKC_PHOSPHO_SITE 275-277; PKC_PHOSPHO_SITE 295-297; MYRISTYL 242-247; MYRISTYL	HIGHMOBLTYIY 249- 268; SER_RICH 209- 302; HIGHMOBLTYIY 222-233; TUBBYPROTEIN 171- 196; TUBBYPROTEIN 285-311; HIGHMOBLTYIY 207- 219; ARG_RICH 207- 339; CYCLIN 13-111;

			315- 320,1.033; 158- 167,1.11; 141- 154,1.056; 90-96,1.106;	30-35; PKC_PHOSPHO_SITE 184-186; CK2_PHOSPHO_SITE 98-101; MYRISTYL 238-243; CK2_PHOSPHO_SITE 121-124; CAMP_PHOSPHO_SITE 232-235; CAMP_PHOSPHO_SITE 313-316; CK2_PHOSPHO_SITE 167-170; MYRISTYL 274-279; AMIDATION 333- 336; PKC_PHOSPHO_SITE 229-231; PKC_PHOSPHO_SITE 11-13; MYRISTYL 143-148; ASN GLYCOSYLATION 53-56; PKC_PHOSPHO_SITE 167-169; MYRISTYL 202-207; TYR_PHOSPHO_SITE 214-220; PKC_PHOSPHO_SITE 55-57; CK2_PHOSPHO_SITE 252-255;	
DEX0449_04 8.aa.1	Y	2 - i1- 4;tm5- 27;o28- 52;tm53 - 75;i76- 91;	63-85,1.21; 40-61,1.119; 13-28,1.179; 6-11,1.065;	MYRISTYL 59-64; PKC_PHOSPHO_SITE 3-5; PKC_PHOSPHO_SITE 47-49; PKC_PHOSPHO_SITE 26-28;	CD225 1-67;
DEX0449_04 8.orf.1	N	0 - i1- 92;	12-52,1.184; 59-89,1.161;	MYRISTYL 40-45; MYRISTYL 1-6; MYRISTYL 59-64; PKC_PHOSPHO_SITE 55-57; TYR_PHOSPHO_SITE 42-49;	
DEX0449_04 9.aa.1	N	0 - o1- 137;	66-73,1.106; 27-39,1.189; 89-95,1.055; 118- 126,1.081; 128- 134,1.045; 41-61,1.185; 75-83,1.108;	CK2_PHOSPHO_SITE 107-110; PKC_PHOSPHO_SITE 29-31; PKC_PHOSPHO_SITE 64-66; MYRISTYL 39-44; CK2_PHOSPHO_SITE 6-9; CK2_PHOSPHO_SITE 62-65; PKC PHOSPHO SITE	Ribophorin_I 2-137;

				24-26; PKC_PHOSPHO_SITE 128-130; TYR_PHOSPHO_SITE 108-115; PKC_PHOSPHO_SITE 87-89;	
DEX0449_04 9.orf.1	N	0 - 01- 148;		MYRISTYL 50-55; PKC_PHOSPHO_SITE 98-100; PKC_PHOSPHO_SITE 35-37; CK2_PHOSPHO_SITE 118-121; PKC_PHOSPHO_SITE 139-141; CK2_PHOSPHO_SITE 22-25; TYR_PHOSPHO_SITE 119-126; CK2_PHOSPHO_SITE 17-20; TYR_PHOSPHO_SITE 16-24; CK2_PHOSPHO_SITE 73-76; PKC_PHOSPHO_SITE 40-42; PKC_PHOSPHO_SITE 75-77;	Ribophorin_I 3-148;
DEX0449_05 0.aa.1	Y	0 - 01- 479;	466- 476,1.085; 230- 237,1.111; 394-403,1.1; 126- 146,1.185; 5-17,1.227; 151- 158,1.106; 253- 261,1.094; 405- 448,1.182; 160- 168,1.108; 300- 311,1.133; 174- 180,1.055; 67-76,1.118; 328- 376,1.147; 37-59,1.187; 203- 211,1.081; 112- 124,1.189; 280-	PKC_PHOSPHO_SITE 369-371; CK2_PHOSPHO_SITE 470-473; MYRISTYL 124-129; MYRISTYL 298-303; CK2_PHOSPHO_SITE 453-456; MYRISTYL 59-64; TYR_PHOSPHO_SITE 391-398; CK2_PHOSPHO_SITE 399-402; PKC_PHOSPHO_SITE 63-65; AMIDATION 462-465; PKC_PHOSPHO_SITE 276-278; MYRISTYL 22-27; PKC_PHOSPHO_SITE 213-215; PKC_PHOSPHO_SITE 114-116; CK2_PHOSPHO_SITE 437-440; MYRISTYL 60-65; PKC_PHOSPHO_SITE 149-151; CK2 PHOSPHO SITE	Ribophorin_I 30- 437;

			294,1.102; 26-34,1.065; 78-89,1.153;	192-195; ASN_GLYCOSYLATION 299-302; CAMP_PHOSPHO_SITE 450-453; CK2_PHOSPHO_SITE 147-150; PKC_PHOSPHO_SITE 109-111; CK2_PHOSPHO_SITE 385-388; CK2_PHOSPHO_SITE 262-265; PKC_PHOSPHO_SITE 462-464; PKC_PHOSPHO_SITE 279-281; MYRISTYL 324-329; TYR_PHOSPHO_SITE 193-200; PKC_PHOSPHO_SITE 172-174;	
DEX0449_05 0.aa.2	Y	0 - 01- 361;	160- 168,1.108; 67-76,1.118; 37-59,1.187; 230- 237,1.111; 280- 294,1.102; 112- 124,1.189; 328- 354,1.108; 26-34,1.065; 5-17,1.227; 300- 311,1.133; 203- 211,1.081; 151- 158,1.106; 126- 146,1.185; 253- 261,1.094; 174- 180,1.055; 78-89,1.153;	MYRISTYL 124-129; TYR_PHOSPHO_SITE 193-200; PKC_PHOSPHO_SITE 149-151; MYRISTYL 324-329; MYRISTYL 22-27; PKC_PHOSPHO_SITE 172-174; CK2_PHOSPHO_SITE 147-150; MYRISTYL 298-303; PKC_PHOSPHO_SITE 114-116; PKC_PHOSPHO_SITE 276-278; CK2_PHOSPHO_SITE 192-195; PKC_PHOSPHO_SITE 63-65; PKC_PHOSPHO_SITE 213-215; ASN_GLYCOSYLATION 299-302; PKC_PHOSPHO_SITE 109-111; MYRISTYL 60-65; PKC_PHOSPHO_SITE 279-281; MYRISTYL 59-64; CK2_PHOSPHO_SITE 262-265;	Ribophorin_I 30- 361;
DEX0449_05 0.aa.3	Y	0 - 01- 420;	395- 406,1.196; 26-34,1.065; 37-59,1.187; 5-17.1.227:	MYRISTYL 124-129; MYRISTYL 22-27; PKC_PHOSPHO_SITE 410-412; MYRISTYL 377-382:	Ribophorin_I 30- 412;

			67-76,1.118; 174- 180,1.055; 126- 146,1.185; 410- 417,1.041; 230- 237,1.111; 151- 158,1.106; 300- 311,1.133; 203- 211,1.081; 160- 168,1.108; 78-89,1.153; 328- 376,1.147; 253- 261,1.094; 280- 294,1.102; 112- 124,1.189; 383- 393,1.162;	PKC_PHOSPHO_SITE 172-174; CK2_PHOSPHO_SITE 192-195; PKC_PHOSPHO_SITE 109-111; MYRISTYL 394-399; MYRISTYL 59-64; PKC_PHOSPHO_SITE 114-116; CK2_PHOSPHO_SITE 262-265; TYR_PHOSPHO_SITE 193-200; PKC_PHOSPHO_SITE 149-151; ASN_GLYCOSYLATION 299-302; MYRISTYL 60-65; PKC_PHOSPHO_SITE 276-278; CK2_PHOSPHO_SITE 147-150; MYRISTYL 324-329; PKC_PHOSPHO_SITE 279-281; PKC_PHOSPHO_SITE 63-65; MYRISTYL 298-303; PKC_PHOSPHO_SITE 213-215; PKC_PHOSPHO_SITE 369-371;	
DEX0449_05 1.aa.1	N	0 - 01- 128;	8-13,1.039; 68-76,1.059; 49-64,1.126; 104- 117,1.058;	MYRISTYL 33-38; MYRISTYL 116-121; PKC_PHOSPHO_SITE 106-108; PKC_PHOSPHO_SITE 24-26; MYRISTYL 44-49; MYRISTYL 91-96; CK2_PHOSPHO_SITE 95-98; MYRISTYL 14-19; MYRISTYL 112-117;	
DEX0449_05 2.aa.1	N	0 - 01- 315;	170- 187,1.133; 93- 115,1.194; 256- 263,1.083; 35-54,1.172; 58-81,1.224; 294- 309,1.089; 267- 273,1.122; 17-30,1.123; 199-	CK2_PHOSPHO_SITE 160-163; MYRISTYL 81-86; ASN_GLYCOSYLATION 155-158; MYRISTYL 136-141; CK2_PHOSPHO_SITE 245-248; MYRISTYL 115-120; PKC_PHOSPHO_SITE 21-23; CK2_PHOSPHO_SITE 15-18; MYRISTYL 5-10;	ank 124-156; ANK 124-153; ank 57-89; ank 91-123; ANKYRIN 137-149; ANKYRIN 125-137; ANK_REPEAT_REGION 21- 179; ANK 21-50; ank 21-55; ANK_REPEAT_3 124-156; ANK_REPEAT_2 91- 123; ANK_REPEAT_1 57-89; ank 157-190; ANK 57-86; ANK 91- 120;

185

			208,1.083; 83-89,1.046; 159- 165,1.053; 137- 154,1.191; 233- 241,1.127;	ASN_GLYCOSYLATION 85-88; PKC_PHOSPHO_SITE 88-90; MYRISTYL 196-201; CK2_PHOSPHO_SITE 241-244; MYRISTYL 192-197; MYRISTYL 158-163; MYRISTYL 223-228; MYRISTYL 9-14; PKC_PHOSPHO_SITE 210-212;	
DEX0449_05 3.aa.1	N	0 - 01- 142;	40-48,1.074; 64-94,1.145; 108- 117,1.083;	CAMP_PHOSPHO_SITE 17-20; PKC_PHOSPHO_SITE 20-22; PKC_PHOSPHO_SITE 132-134; MYRISTYL 38-43; PKC_PHOSPHO_SITE 137-139; PKC_PHOSPHO_SITE 10-12; CAMP_PHOSPHO_SITE 135-138; PKC_PHOSPHO_SITE 21-23; MYRISTYL 116-121; MYRISTYL 2-7; CK2_PHOSPHO_SITE 93-96; CAMP_PHOSPHO_SITE 134-137; AMIDATION 15-18; AMIDATION 120- 123;	
DEX0449_05 3.aa.2	N	0 - 01- 72;	14-32,1.094; 40-46,1.064;	CAMP_PHOSPHO_SITE 65-68; CK2_PHOSPHO_SITE 38-41; CAMP_PHOSPHO_SITE 64-67; MYRISTYL 46-51; PKC_PHOSPHO_SITE 62-64; PKC_PHOSPHO_SITE 67-69; AMIDATION 50-53; MYRISTYL 2-7;	
DEX0449_05 3.orf.2	N	0 - 01- 122;	67- 108,1.199; 59-65,1.06; 11-47,1.156;	CK2_PHOSPHO_SITE 1-4; AMIDATION 115-118; ASN_GLYCOSYLATION 100-103; MYRISTYL 13-18; PKC_PHOSPHO_SITE 59-61; MYRISTYL 27-32;	

				ASN_GLYCOSYLATION 107-110; PKC_PHOSPHO_SITE 95-97; MYRISTYL 53-58;	
DEX0449_05 4.aa.1	N	0 - 01- 42;	5-11,1.101; 29-39,1.142;	MYRISTYL 15-20; CK2_PHOSPHO_SITE 23-26; MYRISTYL 11-16; MYRISTYL 20-25; MYRISTYL 6-11; MYRISTYL 16-21;	
DEX0449_05 4.orf.1	N	0 - 01- 260;	54-80,1.148; 187- 194,1.044; 197- 207,1.183; 86-96,1.133; 34-52,1.074; 118- 126,1.089; 14-24,1.051;	CK2_PHOSPHO_SITE 96-99; CK2_PHOSPHO_SITE 145-148; PKC_PHOSPHO_SITE 145-147; CK2_PHOSPHO_SITE 91-94; PKC_PHOSPHO_SITE 135-137; ASN_GLYCOSYLATION 129-132; PKC_PHOSPHO_SITE 161-163; CAMP_PHOSPHO_SITE 151-154; PKC_PHOSPHO_SITE 150-152;	GLU_RICH 211-258; ASP_RICH 178-260; NAP 12-209;
DEX0449_05 4.orf.2	N	0 - 01- 285;	111- 121,1.133; 212- 219,1.044; 79- 105,1.148; 4-33,1.166; 39-45,1.051; 222- 232,1.183; 59-77,1.074; 143- 151,1.089;	PKC_PHOSPHO_SITE 170-172; PKC_PHOSPHO_SITE 160-162; PKC_PHOSPHO_SITE 186-188; ASN_GLYCOSYLATION 154-157; CK2_PHOSPHO_SITE 121-124; CK2_PHOSPHO_SITE 1-4; CK2_PHOSPHO_SITE 170-173; CAMP_PHOSPHO_SITE 176-179; PKC_PHOSPHO_SITE 175-177; CK2_PHOSPHO_SITE 116-119;	ASP_RICH 203-285; NAP 37-234; GLU_RICH 236-283;
DEX0449_05 5.aa.1	N	0 - 01- 600;	462- 469,1.068; 36-47,1.086; 96- 102,1.051; 494- 506,1.08; 263-	CK2_PHOSPHO_SITE 309-312; CK2_PHOSPHO_SITE 245-248; ASN_GLYCOSYLATION 112-115; MYRISTYL 515-520; CK2 PHOSPHO SITE	IF 368-376; IF_tail 422-559; PRENYLATION 597- 600; filament 25- 381;

			272,1.055; 170- 176,1.091; 248- 261,1.123; 541- 547,1.129; 189- 198,1.061; 427- 434,1.067; 219- 227,1.108; 508- 530,1.101; 13-19,1.061; 295- 309,1.102; 83-89,1.053; 52-58,1.102; 154- 166,1.071; 69-79,1.095; 348- 365,1.142; 484- 490,1.068; 441- 458,1.086; 382- 397,1.091; 122- 148,1.132;	59-62; PKC_PHOSPHO_SITE 408-410; ASN_GLYCOSYLATION 470-473; MYRISTYL 433-438; TYR_PHOSPHO_SITE 291-299; MYRISTYL 435-440; PKC_PHOSPHO_SITE 226-228; PKC_PHOSPHO_SITE 65-67; PKC_PHOSPHO_SITE 472-474; PKC_PHOSPHO_SITE 498-500; PKC_PHOSPHO_SITE 148-150; PKC_PHOSPHO_SITE 114-116; TYR_PHOSPHO_SITE 255-262; CK2_PHOSPHO_SITE 204-207; PKC_PHOSPHO_SITE 90-92; PKC_PHOSPHO_SITE 414-416; MYRISTYL 403-408; MYRISTYL 426-431; PKC_PHOSPHO_SITE 389-391; PKC_PHOSPHO_SITE 556-558; CAMP_PHOSPHO_SITE 558-561; MYRISTYL 506-511; PKC_PHOSPHO_SITE 59-61; CK2_PHOSPHO_SITE 65-68; PKC_PHOSPHO_SITE 268-270; CK2_PHOSPHO_SITE 455-458; CK2_PHOSPHO_SITE 194-197; PKC_PHOSPHO_SITE 213-215; CK2_PHOSPHO_SITE 139-142; MYRISTYL 535-540; CK2_PHOSPHO_SITE 472-475; CK2_PHOSPHO_SITE 90-93; PKC_PHOSPHO_SITE 539-541;	
--	--	--	--	--	--

				PKC_PHOSPHO_SITE 593-595; MYRISTYL 437-442;	
DEX0449_05 5.orf.1	N	0 - ol- 620;	447- 454,1.067; 368- 385,1.142; 315- 329,1.102; 142- 168,1.132; 33-39,1.061; 116- 122,1.051; 103- 109,1.053; 402- 417,1.091; 504- 510,1.068; 461- 478,1.086; 209- 218,1.061; 72-78,1.102; 482- 489,1.068; 89-99,1.095; 528- 550,1.101; 174- 186,1.071; 56-67,1.086; 190- 196,1.091; 239- 247,1.108; 268- 281,1.123; 514- 526,1.08; 283- 292,1.055; 561- 567,1.129; 22-27,1.014;	PKC_PHOSPHO_SITE 134-136; PKC_PHOSPHO_SITE 409-411; MYRISTYL 453-458; MYRISTYL 446-451; PKC_PHOSPHO_SITE 576-578; PKC_PHOSPHO_SITE 246-248; CK2_PHOSPHO_SITE 85-88; PKC_PHOSPHO_SITE 288-290; CK2_PHOSPHO_SITE 214-217; MYRISTYL 555-560; CK2_PHOSPHO_SITE 492-495; PKC_PHOSPHO_SITE 168-170; PKC_PHOSPHO_SITE 613-615; PKC_PHOSPHO_SITE 428-430; TYR_PHOSPHO_SITE 275-282; PKC_PHOSPHO_SITE 434-436; CK2_PHOSPHO_SITE 159-162; PKC_PHOSPHO_SITE 518-520; TYR_PHOSPHO_SITE 311-319; CK2_PHOSPHO_SITE 475-478; CAMP_PHOSPHO_SITE 578-581; PKC_PHOSPHO_SITE 492-494; CK2_PHOSPHO_SITE 224-227; ASN_GLYCOSYLATION 490-493; PKC_PHOSPHO_SITE 233-235; ASN_GLYCOSYLATION 132-135; CK2_PHOSPHO_SITE 265-268; MYRISTYL 423-428; CK2_PHOSPHO_SITE 329-332; PKC_PHOSPHO_SITE 79-81;	PRENYLATION 617- 620; IF_tail 442- 579; IF 388-396; filament 45-401;

				CK2_PHOSPHO_SITE 79-82; MYRISTYL 455-460; PKC_PHOSPHO_SITE 110-112; MYRISTYL 535-540; AMIDATION 6-9; PKC_PHOSPHO_SITE 559-561; PKC_PHOSPHO_SITE 85-87; CK2_PHOSPHO_SITE 110-113; MYRISTYL 526-531; MYRISTYL 457-462;	
DEX0449_05 5.aa.2	N	0 - ol- 237;	174- 186,1.071; 33-39,1.061; 199- 207,1.122; 218- 234,1.134; 103- 109,1.053; 56-67,1.086; 142- 168,1.132; 89-99,1.095; 22-27,1.014; 72-78,1.102; 190- 196,1.091; 116- 122,1.051;	CK2_PHOSPHO_SITE 79-82; PKC_PHOSPHO_SITE 110-112; CK2_PHOSPHO_SITE 85-88; ASN_GLYCOSYLATION 132-135; AMIDATION 6-9; PKC_PHOSPHO_SITE 79-81; PKC_PHOSPHO_SITE 85-87; MYRISTYL 201-206; PKC_PHOSPHO_SITE 205-207; PKC_PHOSPHO_SITE 168-170; CK2_PHOSPHO_SITE 110-113; CK2_PHOSPHO_SITE 159-162; PKC_PHOSPHO_SITE 134-136; CK2_PHOSPHO_SITE 210-213;	filament 45-232;
DEX0449_05 5.orf.3	N	0 - ol- 620;	482- 489,1.068; 268- 281,1.123; 22-27,1.014; 561- 567,1.129; 56-67,1.086; 174- 186,1.071; 368- 385,1.142; 528- 550,1.101; 447- 454,1.067; 72-78,1.102; 283-	CAMP_PHOSPHO_SITE 578-581; TYR_PHOSPHO_SITE 275-282; CK2_PHOSPHO_SITE 329-332; AMIDATION 6-9; CK2_PHOSPHO_SITE 475-478; PKC_PHOSPHO_SITE 613-615; PKC_PHOSPHO_SITE 85-87; MYRISTYL 446-451; CK2_PHOSPHO_SITE 110-113; MYRISTYL 555-560; CK2 PHOSPHO SITE	PRENYLATION 617- 620; IF 388-396; filament 45-401; IF_tail 442-579;

			292,1.055; 209- 218,1.061; 89-99,1.095; 514- 526,1.08; 33-39,1.061; 461- 478,1.086; 504- 510,1.068; 315- 329,1.102; 103- 109,1.053; 239- 247,1.108; 402- 417,1.091; 190- 196,1.091; 142- 168,1.132; 116- 122,1.051;	79-82; CK2_PHOSPHO_SITE 85-88; PKC_PHOSPHO_SITE 559-561; TYR_PHOSPHO_SITE 311-319; PKC_PHOSPHO_SITE 79-81; PKC_PHOSPHO_SITE 168-170; MYRISTYL 457-462; CK2_PHOSPHO_SITE 159-162; PKC_PHOSPHO_SITE 134-136; MYRISTYL 455-460; MYRISTYL 453-458; ASN_GLYCOSYLATION 132-135; CK2_PHOSPHO_SITE 492-495; ASN_GLYCOSYLATION 490-493; PKC_PHOSPHO_SITE 428-430; CK2_PHOSPHO_SITE 214-217; MYRISTYL 526-531; PKC_PHOSPHO_SITE 576-578; PKC_PHOSPHO_SITE 409-411; PKC_PHOSPHO_SITE 288-290; PKC_PHOSPHO_SITE 110-112; PKC_PHOSPHO_SITE 434-436; MYRISTYL 535-540; PKC_PHOSPHO_SITE 233-235; MYRISTYL 423-428; PKC_PHOSPHO_SITE 492-494; CK2_PHOSPHO_SITE 265-268; PKC_PHOSPHO_SITE 518-520; CK2_PHOSPHO_SITE 224-227; PKC_PHOSPHO_SITE 246-248;	
--	--	--	---	--	--

Example 1b: Sequence Alignment Support

Alignments between previously identified sequences and splice variant sequences are performed to confirm unique portions of splice variant nucleic acid and amino acid sequences. The alignments are done using the Needle program in the European Molecular Biology Open Software Suite (EMBOSS) version 2.2.0 available at www.emboss.org from EMBnet (<http://www.embnnet.org>). Default settings are used unless otherwise noted. The Needle program in EMBOSS implements the Needleman-Wunsch algorithm. Needleman, S. B., Wunsch, C. D., *J. Mol. Biol.* 48:443-453 (1970).

It is well known to those skilled in the art that implication of alignment algorithms by various programs may result in minor changes in the generated output. These changes include but are not limited to: alignment scores (percent identity, similarity, and gap), display of nonaligned flanking sequence regions, and number assignment to residues. These minor changes in the output of an alignment do not alter the physical characteristics of the sequences or the differences between the sequences, e.g. regions of homology, insertions, or deletions.

15 **Example 1c: RT-PCR Analysis**

To detect the presence and tissue distribution of a particular splice variant Reverse Transcription-Polymerase Chain Reaction (RT-PCR) is performed using cDNA generated from a panel of tissue RNAs. See, e.g., Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and; Kawasaki ES *et al.*, *PNAS* 85(15):5698 (1988). Total RNA is extracted from a variety of tissues and first strand cDNA is prepared with reverse transcriptase (RT). Each panel includes 23 cDNAs from five cancer types (lung, ovary, breast, colon, and prostate) and normal samples of testis, placenta and fetal brain. Each cancer set is composed of three cancer cDNAs from different donors and one normal pooled sample. Using a standard enzyme kit from BD Bioscience Clontech (Mountain View, CA), the target transcript is detected with sequence-specific primers designed to only amplify the particular splice variant. The PCR reaction is run on the GeneAmp PCR system 9700 (Applied Biosystem, Foster City, CA) thermocycler under optimal conditions. One of ordinary skill can design appropriate primers and determine optimal conditions. The amplified product is resolved on an agarose gel to detect a band of equivalent size to the predicted RT-PCR product. A band indicated the presence of the splice variant in a sample. The relation of the amplified product to the splice variant was subsequently confirmed by DNA sequencing.

After subcloning, all positively screened clones are sequence verified. The DNA sequence verification results show the splice variant contains the predicted sequence differences in comparison with the reference sequence.

Results for RT-PCR analysis include the sequence DEX ID, Lead Name, Cancer
5 Tissue(s) the transcript was detected in, Normal Tissue(s) the transcript was detected in, the predicted length of the RT-PCR product, and the Confirmed Length of the RT-PCR product.

RT-PCR results confirm the presence SEQ ID NO: 1-112 in biologic samples and distinguish between related transcripts.

10 **Example 1d: Secretion Assay**

To determine if a protein encoded by a splice variant is secreted from cells a secretion assay is preformed. A pcDNA3.1 clone containing the gene transcript which encodes the variant protein is transfected into 293T cells using the Superfect transfection reagent (Qiagen, Valencia CA). Transfected cells are incubated for 28 hours before the
15 media is collected and immediately spun down to remove any detached cells. The adherent cells are solubilized with lysis buffer (1% NP40, 10mM sodium phosphate pH7.0, and 0.15M NaCl). The lysed cells are collected and spun down and the supernatant extracted as cell lysate. Western immunoblot is carried out in the following manner: 15µl of the cell lysate and media are run on 4-12% NuPage Bis-Tris gel
20 (Invitrogen, Carlsbad CA), and blotted onto a PVDF membrane (Invitrogen, Carlsbad CA). The blot is incubated with a polyclonal primary antibody which binds to the variant protein (Imgenex, San Diego CA) and polyclonal goat anti-rabbit-peroxidase secondary antibody (Sigma-Aldrich, St. Louis MO). The blot is developed with the ECL Plus chemiluminescent detection reagent (Amersham BioSciences, Piscataway NJ).

25 Secretion assay results are indicative of SEQ ID NO: 113-259 being a diagnostic marker and/or therapeutic target for cancer.

Example 2a: Gene Expression Analysis

Custom Microarray Experiment - Cancer

Custom oligonucleotide microarrays were provided by Agilent Technologies, Inc.
30 (Palo Alto, CA). The microarrays were fabricated by Agilent using their technology for the *in-situ* synthesis of 60mer oligonucleotides (Hughes, et al. 2001, Nature Biotechnology 19:342-347). The 60mer microarray probes were designed by Agilent, from gene

sequences provided by diaDexus, using Agilent proprietary algorithms. Whenever possible two different 60mers were designed for each gene of interest.

All microarray experiments were two-color experiments and were preformed using Agilent-recommended protocols and reagents. Briefly, each microarray was hybridized
5 with cRNAs synthesized from RNA (total RNA for ovarian and prostate, polyA+ RNA for lung, breast and colon samples), isolated from cancer and normal tissues, labeled with fluorescent dyes Cyanine3 (Cy3) or Cyanine5 (Cy5) (NEN Life Science Products, Inc., Boston, MA) using a linear amplification method (Agilent). In each experiment the experimental sample was RNA isolated from cancer tissue from a single individual and
10 the reference sample was a pool of RNA isolated from normal tissues of the same organ as the cancerous tissue (*i.e.* normal ovarian tissue in experiments with ovarian cancer samples). Hybridizations were carried out at 60°C, overnight using Agilent *in-situ* hybridization buffer. Following washing, arrays were scanned with a GenePix 4000B Microarray Scanner (Axon Instruments, Inc., Union City, CA). The resulting images were
15 analyzed with GenePix Pro 3.0 Microarray Acquisition and Analysis Software (Axon).

Data normalization and expression profiling were done with Expressionist software from GeneData Inc. (Daly City, CA/Basel, Switzerland). Gene expression analysis was performed using only experiments that met certain quality criteria. The quality criteria that experiments must meet are a combination of evaluations performed by
20 the Expressionist software and evaluations performed manually using raw and normalized data. To evaluate raw data quality, detection limits (the mean signal for a replicated negative control + 2 Standard Deviations (SD)) for each channel were calculated. The detection limit is a measure of non-specific hybridization. Acceptable detection limits were defined for each dye (<80 for Cy5 and <150 for Cy3). Arrays with poor detection
25 limits in one or both channels were not analyzed and the experiments were repeated. To evaluate normalized data quality, positive control elements included in the array were utilized. These array features should have a mean ratio of 1 (no differential expression). If these features have a mean ratio of greater than 1.5-fold up or down, the experiments were not analyzed further and were repeated. In addition to traditional scatter plots
30 demonstrating the distribution of signal in each experiment, the Expressionist software also has minimum thresholding criteria that employ user defined parameters to identify quality data. These thresholds include two distinct quality measurements: 1) minimum area percentage, which is a measure of the integrity of each spot and 2) signal to noise

ratio, which ensures that the signal being measured is significantly above any background (nonspecific) signal present. Only those features that met the threshold criteria were included in the filtering and analyses carried out by Expressionist. The thresholding settings employed require a minimum area percentage of 60% [(% pixels > background + 2SD)-(% pixels saturated)], and a minimum signal to noise ratio of 2.0 in both channels. By these criteria, very low expressors, saturated features and spots with abnormally high local background were not included in analysis.

Relative expression data was collected from Expressionist based on filtering and clustering analyses. Up-regulated genes were identified using criteria for the percentage of experiments in which the gene is up-regulated by at least 2-fold. In general, up-regulation in ~30% of samples tested was used as a cutoff for filtering.

Two microarray experiments were performed for each normal and cancer tissue pair. The tissue specific Array Chip for each cancer tissue is a unique microarray specific to that tissue and cancer. The Multi-Cancer Array Chip is a universal microarray that was hybridized with samples from each of the cancers (ovarian, breast, colon, lung, and prostate). See the description below for the experiments specific to the different cancers.

Microarray Experiments and Data Tables

COLON CANCER CHIPS

For colon cancer two different chip designs were evaluated with overlapping sets of a total of 38 samples, comparing the expression patterns of colon cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 7 normal colon tissues. For the Colon Array Chip all 38 samples (23 Ascending colon carcinomas and 15 Rectosigmoidal carcinomas including: 5 stage I cancers, 15 stage II cancers, 15 stage III and 2 stage IV cancers, as well as 28 Grade 1/2 and 10 Grade 3 cancers) were analyzed. The histopathologic grades for cancer are classified as follows: GX, cannot be assessed; G1, well differentiated; G2, Moderately differentiated; G3, poorly differentiated; and G4, undifferentiated. AJCC Cancer Staging Handbook, 5th Edition, 1998, page 9. For the Colon Array Chip analysis, samples were further divided into groups based on the expression pattern of the known colon cancer associated gene Thymidilate Synthase (TS) (13 TS up 25 TS not up). The association of TS with advanced colorectal cancer is well documented. Paradiso *et al.*, *Br J Cancer* 82(3):560-7 (2000); Etienne *et al.*, *J Clin Oncol.* 20(12):2832-43 (2002); Aschele *et al.* *Clin Cancer Res.* 6(12):4797-802 (2000).

For the Multi-Cancer Array Chip a subset of 27 of these samples (14 Ascending colon carcinomas and 13 Rectosigmoidal carcinomas including: 3 stage I cancers, 9 stage II cancers, 13 stage III and 2 stage IV cancers) were assessed.

- The results for the statistically significant up-regulated genes on the Colon Array Chip are shown in Tables 1 and 2. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 3.

The first two columns of each table contain information about the sequence itself (Seq ID, Oligo Name), the next columns show the results obtained for all ("ALL") the colon samples, ascending colon carcinomas ("ASC"), Rectosigmoidal carcinomas ("RS"), cancers corresponding to stages I and II ("ST1,2"), stages III and IV ("ST3,4"), grades 1 and 2 ("GR1,2"), grade 3 ("GR3"), cancers exhibiting up-regulation of the TS gene ("TSup") or those not exhibiting up-regulation of the TS gene ("NOT TSup"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed n=38 for the Colon Array Chip (n=27 for the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 1.

DEX ID	Oligo Name	Cln ALL %up n=38	Cln ALL % valid up n=38	Cln ASC %up n=23	Cln ASC % valid up n=23	Cln RS %up n=15	Cln RS % valid up n=15	Cln ST1,2 %up n=20	Cln ST1,2 % valid up n=20	Cln ST3,4 %up n=18	Cln ST3,4 % valid up n=18
DEX0449_001.nt.1	10296.0	44.7	44.7	60.9	60.9	20.0	20.0	45.0	45.0	44.4	44.4
DEX0449_001.nt.1	10297.0	63.2	63.2	73.9	73.9	46.7	46.7	65.0	65.0	61.1	61.1
DEX0449_044.nt.1	33264.0	5.3	8.0	0.0	0.0	13.3	25.0	0.0	0.0	11.1	15.4
DEX0449_044.nt.1	33265.0	10.5	11.8	4.3	5.0	20.0	21.4	5.0	5.6	16.7	18.8
DEX0449_044.nt.2	33264.0	5.3	8.0	0.0	0.0	13.3	25.0	0.0	0.0	11.1	15.4
DEX0449_044.nt.2	33265.0	10.5	11.8	4.3	5.0	20.0	21.4	5.0	5.6	16.7	18.8
DEX0449_044.nt.3	33264.0	5.3	8.0	0.0	0.0	13.3	25.0	0.0	0.0	11.1	15.4
DEX0449_044.nt.3	33265.0	10.5	11.8	4.3	5.0	20.0	21.4	5.0	5.6	16.7	18.8
DEX0449_044.nt.4	33264.0	5.3	8.0	0.0	0.0	13.3	25.0	0.0	0.0	11.1	15.4
DEX0449_044.nt.4	33265.0	10.5	11.8	4.3	5.0	20.0	21.4	5.0	5.6	16.7	18.8
DEX0449_046.nt.1	33720.0	36.8	36.8	43.5	43.5	26.7	26.7	35.0	35.0	38.9	38.9

DEX0449_047.nt.1	12809.0	15.8	15.8	13.0	13.0	20.0	20.0	5.0	5.0	27.8	27.8
DEX0449_047.nt.1	12810.0	15.8	15.8	13.0	13.0	20.0	20.0	5.0	5.0	27.8	27.8
DEX0449_047.nt.2	12809.0	15.8	15.8	13.0	13.0	20.0	20.0	5.0	5.0	27.8	27.8
DEX0449_047.nt.2	12810.0	15.8	15.8	13.0	13.0	20.0	20.0	5.0	5.0	27.8	27.8
DEX0449_048.nt.1	28637.0	73.7	73.7	82.6	82.6	60.0	60.0	70.0	70.0	77.8	77.8
DEX0449_048.nt.1	28638.0	65.8	65.8	73.9	73.9	53.3	53.3	70.0	70.0	61.1	61.1
DEX0449_049.nt.1	10224.0	13.2	13.9	17.4	18.2	6.7	7.1	10.0	10.0	16.7	18.8
DEX0449_049.nt.1	10225.0	7.9	8.8	8.7	9.5	6.7	7.7	0.0	0.0	16.7	20.0
DEX0449_051.nt.1	28403.0	10.5	10.5	17.4	17.4	0.0	0.0	10.0	10.0	11.1	11.1
DEX0449_052.nt.1	16385.0	21.1	21.1	21.7	21.7	20.0	20.0	20.0	20.0	22.2	22.2
DEX0449_053.nt.1	17654.0	5.3	5.6	4.3	4.5	6.7	7.1	0.0	0.0	11.1	11.1
DEX0449_053.nt.1	17655.0	13.2	13.9	17.4	19.0	6.7	6.7	10.0	10.5	16.7	17.6
DEX0449_053.nt.2	17654.0	5.3	5.6	4.3	4.5	6.7	7.1	0.0	0.0	11.1	11.1
DEX0449_053.nt.2	17655.0	13.2	13.9	17.4	19.0	6.7	6.7	10.0	10.5	16.7	17.6
DEX0449_054.nt.1	37943.0	52.6	52.6	56.5	56.5	46.7	46.7	50.0	50.0	55.6	55.6
DEX0449_054.nt.1	37944.0	50.0	51.4	56.5	56.5	40.0	42.9	50.0	52.6	50.0	50.0
DEX0449_054.nt.2	37943.0	52.6	52.6	56.5	56.5	46.7	46.7	50.0	50.0	55.6	55.6
DEX0449_054.nt.2	37944.0	50.0	51.4	56.5	56.5	40.0	42.9	50.0	52.6	50.0	50.0
DEX0449_055.nt.1	30951.0	65.8	65.8	69.6	69.6	60.0	60.0	65.0	65.0	66.7	66.7
DEX0449_055.nt.1	30956.0	55.3	56.8	69.6	72.7	33.3	33.3	50.0	52.6	61.1	61.1
DEX0449_055.nt.1	31201.0	65.8	65.8	69.6	69.6	60.0	60.0	70.0	70.0	61.1	61.1
DEX0449_055.nt.2	30951.0	65.8	65.8	69.6	69.6	60.0	60.0	65.0	65.0	66.7	66.7
DEX0449_055.nt.2	30956.0	55.3	56.8	69.6	72.7	33.3	33.3	50.0	52.6	61.1	61.1
DEX0449_055.nt.2	31201.0	65.8	65.8	69.6	69.6	60.0	60.0	70.0	70.0	61.1	61.1
DEX0449_055.nt.3	30951.0	65.8	65.8	69.6	69.6	60.0	60.0	65.0	65.0	66.7	66.7
DEX0449_055.nt.3	30956.0	55.3	56.8	69.6	72.7	33.3	33.3	50.0	52.6	61.1	61.1
DEX0449_055.nt.3	31201.0	65.8	65.8	69.6	69.6	60.0	60.0	70.0	70.0	61.1	61.1

Table 2.

DEX ID	Oligo Name	Cln GR1,2 %up n=28	Cln GR1,2 %valid up n=28	Cln GR3 %up n=10	Cln GR3 %valid up n=10	Cln TS up %up n=13	Cln TS up %valid up n=13	Cln NOT TS up %up n=25	Cln NOT TS up %valid up n=25
DEX0449 001.nt.1	10296.0	35.7	35.7	70.0	70.0	61.5	61.5	36.0	36.0
DEX0449 001.nt.1	10297.0	53.6	53.6	90.0	90.0	84.6	84.6	52.0	52.0
DEX0449 044.nt.1	33264.0	7.1	11.1	0.0	0.0	0.0	0.0	8.0	11.1
DEX0449 044.nt.1	33265.0	10.7	11.5	10.0	12.5	7.7	11.1	12.0	12.0
DEX0449 044.nt.2	33264.0	7.1	11.1	0.0	0.0	0.0	0.0	8.0	11.1
DEX0449 044.nt.2	33265.0	10.7	11.5	10.0	12.5	7.7	11.1	12.0	12.0
DEX0449 044.nt.3	33264.0	7.1	11.1	0.0	0.0	0.0	0.0	8.0	11.1
DEX0449 044.nt.3	33265.0	10.7	11.5	10.0	12.5	7.7	11.1	12.0	12.0
DEX0449 044.nt.4	33264.0	7.1	11.1	0.0	0.0	0.0	0.0	8.0	11.1
DEX0449 044.nt.4	33265.0	10.7	11.5	10.0	12.5	7.7	11.1	12.0	12.0
DEX0449 046.nt.1	33720.0	28.6	28.6	60.0	60.0	38.5	38.5	36.0	36.0
DEX0449 047.nt.1	12809.0	10.7	10.7	30.0	30.0	15.4	15.4	16.0	16.0
DEX0449 047.nt.1	12810.0	10.7	10.7	30.0	30.0	15.4	15.4	16.0	16.0
DEX0449 047.nt.2	12809.0	10.7	10.7	30.0	30.0	15.4	15.4	16.0	16.0
DEX0449 047.nt.2	12810.0	10.7	10.7	30.0	30.0	15.4	15.4	16.0	16.0
DEX0449 048.nt.1	28637.0	71.4	71.4	80.0	80.0	69.2	69.2	76.0	76.0
DEX0449 048.nt.1	28638.0	64.3	64.3	70.0	70.0	61.5	61.5	68.0	68.0
DEX0449 049.nt.1	10224.0	3.6	3.7	40.0	44.4	30.8	30.8	4.0	4.3
DEX0449 049.nt.1	10225.0	3.6	3.8	20.0	25.0	23.1	23.1	0.0	0.0
DEX0449 051.nt.1	28403.0	3.6	3.6	30.0	30.0	23.1	23.1	4.0	4.0
DEX0449 052.nt.1	16385.0	14.3	14.3	40.0	40.0	30.8	30.8	16.0	16.0
DEX0449 053.nt.1	17654.0	0.0	0.0	20.0	20.0	15.4	16.7	0.0	0.0
DEX0449 053.nt.1	17655.0	7.1	7.4	30.0	33.3	38.5	38.5	0.0	0.0
DEX0449 053.nt.2	17654.0	0.0	0.0	20.0	20.0	15.4	16.7	0.0	0.0
DEX0449 053.nt.2	17655.0	7.1	7.4	30.0	33.3	38.5	38.5	0.0	0.0
DEX0449 054.nt.1	37943.0	46.4	46.4	70.0	70.0	76.9	76.9	40.0	40.0
DEX0449 054.nt.1	37944.0	42.9	44.4	70.0	70.0	76.9	76.9	36.0	37.5
DEX0449 054.nt.2	37943.0	46.4	46.4	70.0	70.0	76.9	76.9	40.0	40.0
DEX0449 054.nt.2	37944.0	42.9	44.4	70.0	70.0	76.9	76.9	36.0	37.5
DEX0449 055.nt.1	30951.0	67.9	67.9	60.0	60.0	100.0	100.0	48.0	48.0
DEX0449 055.nt.1	30956.0	53.6	55.6	60.0	60.0	92.3	92.3	36.0	37.5
DEX0449 055.nt.1	31201.0	67.9	67.9	60.0	60.0	100.0	100.0	48.0	48.0
DEX0449 055.nt.2	30951.0	67.9	67.9	60.0	60.0	100.0	100.0	48.0	48.0
DEX0449 055.nt.2	30956.0	53.6	55.6	60.0	60.0	92.3	92.3	36.0	37.5
DEX0449 055.nt.2	31201.0	67.9	67.9	60.0	60.0	100.0	100.0	48.0	48.0
DEX0449 055.nt.3	30951.0	67.9	67.9	60.0	60.0	100.0	100.0	48.0	48.0
DEX0449 055.nt.3	30956.0	53.6	55.6	60.0	60.0	92.3	92.3	36.0	37.5
DEX0449 055.nt.3	31201.0	67.9	67.9	60.0	60.0	100.0	100.0	48.0	48.0

Table 3.

DEX ID	Oligo Name	Cln Multi-Cancer ALL %up n=27	Cln Multi-Cancer ALL %valid up n=27	Cln Multi-Cancer ASC %up n=14	Cln Multi-Cancer ASC %valid up n=14	Cln Multi-Cancer RS %up n=13	Cln Multi-Cancer RS %valid up n=13
DEX0449 044.nt.1	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92319.0	0.0	0.0	0.0	0.0	0.0	0.0

DEX0449	044.nt.1	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.1	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.1	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.2	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.3	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	044.nt.4	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449	045.nt.1	1045.0	25.9	25.9	14.3	14.3	38.5	38.5

BREAST CANCER CHIPS

- For breast cancer two different chip designs were evaluated with overlapping sets of a total of 36 samples, comparing the expression patterns of breast cancer derived
- 5 polyA⁺ RNA to polyA⁺ RNA isolated from a pool of 10 normal breast tissues. For the Breast Array Chip, all 36 samples (9 stage I cancers, 23 stage II cancers, 4 stage III cancers) were analyzed. These samples also represented 10 Grade 1/2 and 26 Grade 3 cancers. The histopathologic grades for cancer are classified as follows: GX, cannot be assessed; G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated;
- 10 and G4, undifferentiated. AJCC Cancer Staging Handbook, pp. 9, (5th Ed, 1998). Samples were further grouped based on the expression patterns of the known breast cancer associated genes Her2 and ER α (10 HER2 up, 26 HER2 not up, 20 ER up and 16 ER not up) and for the Multi-Cancer Array Chip, a subset of 20 of these samples (9 stage I cancers, 8 stage II cancers, 3 stage III cancers) were assessed.
- 15 No results for the statistically significant up-regulated genes on the Breast Array Chip are shown. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 4. The first two columns of each table

contain information about the sequence itself (Seq ID, Oligo Name), the next columns show the results obtained for all ("ALL") breast cancer samples, cancers corresponding to stage I ("ST1"), stages II and III ("ST2,3"), grades 1 and 2 ("GR1,2"), grade 3 ("GR3"), cancers exhibiting up-regulation of Her2 ("HER2up") or ERα ("ERup") or those not exhibiting up-regulation of Her2 ("NOT HER2up") or ERα ("NOT ERup"). 'up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=36 for Colon Array Chip, n=20 for the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

10 Table 4.

DEX ID	Oligo Name	Mam Multi-Cancer ALL %up n=20	Mam Multi-Cancer ALL %valid up n=20	Mam Multi-Cancer ST1 %up n=9	Mam Multi-Cancer ST1 %valid up n=9	Mam Multi-Cancer ST2,3 %up n=11	Mam Multi-Cancer ST2,3 %valid up n=11
DEX0449 044.nt.1	78951.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.1	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	78952.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.1	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.2	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78952.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.2	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.3	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78952.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.3	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.4	78952.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78952.1	5.0	14.3	0.0	0.0	9.1	16.7
DEX0449 044.nt.4	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	1045.0	0.0	0.0	0.0	0.0	0.0	0.0

LUNG CANCER CHIPS

For lung cancer two different chip designs were evaluated with overlapping sets of a total of 29 samples, comparing the expression patterns of lung cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 12 normal lung tissues. For the Lung Array Chip all 29 samples (15 squamous cell carcinomas and 14 adenocarcinomas including 14 stage I and 15 stage II/III cancers) were analyzed and for the Multi-Cancer Array Chip a subset of 22 of these samples (10 squamous cell carcinomas, 12 adenocarcinomas) were assessed.

The results for the statistically significant up-regulated genes on the Lung Array Chip are shown in Table 5. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 6. The first two columns of each table contain information about the sequence itself (DEX ID, Oligo Name), the next columns show the results obtained for all ("ALL") lung cancer samples, squamous cell carcinomas ("SQ"), adenocarcinomas ("AD"), or cancers corresponding to stage I ("ST1"), or stages II and III ("ST2,3"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=29 for Lung Array Chip, n=22 for Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 5.

DEX ID	Oligo Name	Lng ALL %up n=29	Lng ALL %valid up n=29	Lng SQ %up n=15	Lng SQ %valid up n=15	Lng AD %up n=14	Lng AD %valid up n=14	Lng ST1 %up n=14	Lng ST1 %valid up n=14	Lng ST2,3 %up n=15	Lng ST2,3 %valid up n=15
DEX0449_045.nt.1	1044.0	13.8	13.8	13.3	13.3	14.3	14.3	14.3	14.3	13.3	13.3
DEX0449_052.nt.1	5819.0	6.9	7.1	6.7	6.7	7.1	7.7	7.1	7.7	6.7	6.7
DEX0449_052.nt.1	5820.0	6.9	6.9	6.7	6.7	7.1	7.1	7.1	7.1	6.7	6.7
DEX0449_054.nt.1	6889.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449_054.nt.1	6890.0	6.9	7.1	6.7	7.1	7.1	7.1	14.3	15.4	0.0	0.0
DEX0449_054.nt.2	6889.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449_054.nt.2	6890.0	6.9	7.1	6.7	7.1	7.1	7.1	14.3	15.4	0.0	0.0

Table 6.

DEX ID	Oligo Name	Lng Multi-Cancer ALL %up n=22	Lng Multi-Cancer ALL %valid up n=22	Lng Multi-Cancer SQ %up n=10	Lng Multi-Cancer SQ %valid up n=10	Lng Multi-Cancer AD %up n=12	Lng Multi-Cancer AD %valid up n=12
DEX0449 044.nt.1	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	78952.0	4.5	33.3	10.0	50.0	0.0	0.0
DEX0449 044.nt.1	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78952.0	4.5	33.3	10.0	50.0	0.0	0.0
DEX0449 044.nt.2	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78952.0	4.5	33.3	10.0	50.0	0.0	0.0
DEX0449 044.nt.3	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78952.0	4.5	33.3	10.0	50.0	0.0	0.0
DEX0449 044.nt.4	78952.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	1045.0	4.5	4.5	0.0	0.0	8.3	8.3

OVARIAN CANCER CHIPS

For ovarian cancer two different chip designs were evaluated with overlapping sets of a total of 19 samples, comparing the expression patterns of ovarian cancer derived total RNA to total RNA isolated from a pool of 9 normal ovarian tissues. For the Multi-Cancer Array Chip, all 19 samples (14 invasive carcinomas, 5 low malignant potential samples were analyzed and for the Ovarian Array Chip, a subset of 17 of these samples (13 invasive carcinomas, 4 low malignant potential samples) were assessed.

No results for the statistically significant up-regulated genes on the Ovarian Array Chip are shown. The results for the Multi-Cancer Array Chip are shown in Table 7. The first two columns of each table contain information about the sequence itself (DEX ID,

Oligo Name), the next columns show the results obtained for all ("ALL") ovarian cancer samples, invasive carcinomas ("INV") and low malignant potential ("LMP") samples.

'%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=19 for the Multi-Cancer Array Chip, n=17 for the Ovarian Array Chip),

- 5 '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 7.

DEX ID	Oligo Name	Ovr Multi-Cancer ALL %up n=19	Ovr Multi-Cancer ALL %valid up n=19	Ovr Multi-Cancer INV %up n=14	Ovr Multi-Cancer INV %valid up n=14	Ovr Multi-Cancer LMP %up n=5	Ovr Multi-Cancer LMP %valid up n=5
DEX0449 044.nt.1	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	78952.0	5.3	12.5	7.1	16.7	0.0	0.0
DEX0449 044.nt.1	78952.1	5.3	12.5	7.1	14.3	0.0	0.0
DEX0449 044.nt.1	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78952.0	5.3	12.5	7.1	16.7	0.0	0.0
DEX0449 044.nt.2	78952.1	5.3	12.5	7.1	14.3	0.0	0.0
DEX0449 044.nt.2	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78952.0	5.3	12.5	7.1	16.7	0.0	0.0
DEX0449 044.nt.3	78952.1	5.3	12.5	7.1	14.3	0.0	0.0
DEX0449 044.nt.3	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78952.0	5.3	12.5	7.1	16.7	0.0	0.0
DEX0449 044.nt.4	78952.1	5.3	12.5	7.1	14.3	0.0	0.0
DEX0449 044.nt.4	92319.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92319.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	1045.0	0.0	0.0	0.0	0.0	0.0	0.0

PROSTATE CANCER

10

For prostate cancer three different chip designs were evaluated with overlapping sets of a total of 29 samples, comparing the expression patterns of prostate cancer or

benign disease derived total RNA to total RNA isolated from a pool of 35 normal prostate tissues. For the Prostate 1 Array and Prostate 2 Array Chips all 29 samples (17 prostate cancer samples, 12 non-malignant disease samples) were analyzed. For the Multi-Cancer Array Chip a subset of 28 of these samples (16 prostate cancer samples, 12 non-malignant disease samples) were analyzed.

The results for the statistically significant up-regulated genes on the Prostate1 Array Chip and the Prostate2 Array Chip are shown in Table 8. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 9. The first two columns of each table contain information about the sequence itself (DEX ID, Oligo Name), the next columns show the results obtained for prostate cancer samples ("CAN") or non-malignant disease samples ("DIS"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=29 for the Prostate1 Array Chip and the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 8.

DEX ID	Oligo Name	Pro CAN %up n=17	Pro CAN %valid up n=17	Pro DIS %up n=12	Pro DIS %valid up n=12
DEX0449 045.nt.1	27017.01	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	27017.02	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	28249.01	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	28249.02	0.0	0.0	0.0	0.0
DEX0449 045.nt.3	27017.01	0.0	0.0	0.0	0.0
DEX0449 045.nt.3	27017.02	0.0	0.0	0.0	0.0
DEX0449 047.nt.1	31094.01	0.0	0.0	0.0	0.0
DEX0449 047.nt.1	31094.02	0.0	0.0	0.0	0.0
DEX0449 047.nt.1	31094.03	0.0	0.0	0.0	0.0
DEX0449 047.nt.1	33588.01	5.9	8.3	0.0	0.0
DEX0449 047.nt.1	33588.02	0.0	0.0	0.0	0.0
DEX0449 047.nt.2	31094.01	0.0	0.0	0.0	0.0
DEX0449 047.nt.2	31094.02	0.0	0.0	0.0	0.0
DEX0449 047.nt.2	31094.03	0.0	0.0	0.0	0.0
DEX0449 047.nt.2	33588.01	5.9	8.3	0.0	0.0
DEX0449 047.nt.2	33588.02	0.0	0.0	0.0	0.0
DEX0449 051.nt.1	35746.01	0.0	0.0	0.0	0.0
DEX0449 051.nt.1	35746.02	0.0	0.0	0.0	0.0
DEX0449 051.nt.1	35746.03	0.0	0.0	0.0	0.0
DEX0449 051.nt.1	35788.01	0.0	0.0	0.0	0.0
DEX0449 051.nt.1	35788.02	0.0	0.0	0.0	0.0
DEX0449 051.nt.1	35788.03	0.0	0.0	0.0	0.0

Table 9.

DEX ID	Oligo Name	Pro Multi-Cancer CAN %up n=16	Pro Multi-Cancer CAN %valid up n=16	Pro Multi-Cancer DIS %up n=12	Pro Multi-Cancer DIS %valid up n=12
DEX0449 044.nt.1	78951.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	78952.0	5.9	7.7	0.0	0.0
DEX0449 044.nt.1	78952.1	5.9	7.1	0.0	0.0
DEX0449 044.nt.1	92319.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92319.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.1	92320.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78951.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	78952.0	5.9	7.7	0.0	0.0
DEX0449 044.nt.2	78952.1	5.9	7.1	0.0	0.0
DEX0449 044.nt.2	92319.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92319.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.2	92320.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78951.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	78952.0	5.9	7.7	0.0	0.0
DEX0449 044.nt.3	78952.1	5.9	7.1	0.0	0.0
DEX0449 044.nt.3	92319.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92319.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.3	92320.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78951.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	78952.0	5.9	7.7	0.0	0.0
DEX0449 044.nt.4	78952.1	5.9	7.1	0.0	0.0
DEX0449 044.nt.4	92319.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92319.1	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.0	0.0	0.0	0.0	0.0
DEX0449 044.nt.4	92320.1	0.0	0.0	0.0	0.0
DEX0449 045.nt.1	1045.0	0.0	0.0	0.0	0.0

SEQ ID NO: 1-112 was up-regulated on various tissue microarrays. Accordingly, nucleotide SEQ ID NO: 1-112 or the encoded protein SEQ ID NO: 113-259 may be used as a cancer therapeutic and/or diagnostic target for the tissues in which expression is shown.

The following table lists the location (Oligo Location) where the microarray oligos (Oligo ID) map on the transcripts (DEX ID) of the present invention. Each Oligo ID may have been printed multiple times on a single chip as replicates. The Oligo Name is an exemplary replicate (e.g. 1000.01) for the Oligo ID (e.g. 1000), and data from other replicates (e.g. 1000.02, 1000.03) may be reported. Additionally, the Array (Chip Name) that each oligo and oligo replicates were printed on is included.

DEX NT ID	Oligo ID	Oligo Name	Chip Name	Oligo Location
DEX0449 001.nt.1	10296	10296.0	Colon array	1034-1093
DEX0449 001.nt.1	10297	10297.0	Colon array	287-346
DEX0449 044.nt.1	78952	78952.0	Multi-Cancer array	428-487
DEX0449 044.nt.1	33264	33264.0	Colon array	405-464
DEX0449 044.nt.1	92319	92319.0	Multi-Cancer array	546-605
DEX0449 044.nt.1	78951	78951.0	Multi-Cancer array	425-484
DEX0449 044.nt.1	33265	33265.0	Colon array	356-415
DEX0449 044.nt.1	92320	92320.0	Multi-Cancer array	526-585
DEX0449 044.nt.2	92319	92319.0	Multi-Cancer array	592-651
DEX0449 044.nt.2	92320	92320.0	Multi-Cancer array	572-631
DEX0449 044.nt.2	33264	33264.0	Colon array	451-510
DEX0449 044.nt.2	78952	78952.0	Multi-Cancer array	474-533
DEX0449 044.nt.2	33265	33265.0	Colon array	402-461
DEX0449 044.nt.2	78951	78951.0	Multi-Cancer array	471-530
DEX0449 044.nt.3	92320	92320.0	Multi-Cancer array	359-418
DEX0449 044.nt.3	33265	33265.0	Colon array	189-248
DEX0449 044.nt.3	78952	78952.0	Multi-Cancer array	261-320
DEX0449 044.nt.3	33264	33264.0	Colon array	238-297
DEX0449 044.nt.3	78951	78951.0	Multi-Cancer array	258-317
DEX0449 044.nt.3	92319	92319.0	Multi-Cancer array	379-438
DEX0449 044.nt.4	92320	92320.0	Multi-Cancer array	335-394
DEX0449 044.nt.4	78952	78952.0	Multi-Cancer array	237-296
DEX0449 044.nt.4	33265	33265.0	Colon array	165-224
DEX0449 044.nt.4	92319	92319.0	Multi-Cancer array	355-414
DEX0449 044.nt.4	33264	33264.0	Colon array	214-273
DEX0449 044.nt.4	78951	78951.0	Multi-Cancer array	234-293
DEX0449 045.nt.1	1044	1044.0	Lung array	175-234
DEX0449 045.nt.1	27017	27017.01	Prostate1 array	1043-1102
DEX0449 045.nt.1	1045	1045.0	Multi-Cancer array	165-224
DEX0449 045.nt.1	28249	28249.02	Prostate1 array	323-382
DEX0449 045.nt.2	1044	1044.0	Lung array	175-234
DEX0449 045.nt.2	1045	1045.0	Multi-Cancer array	165-224
DEX0449 045.nt.3	1044	1044.0	Lung array	160-219
DEX0449 045.nt.3	27017	27017.01	Prostate1 array	1613-1672
DEX0449 045.nt.3	1045	1045.0	Multi-Cancer array	150-209
DEX0449 046.nt.1	33720	33720.0	Colon array	1037-1096
DEX0449 047.nt.1	12809	12809.0	Colon array	2811-2870
DEX0449 047.nt.1	33588	33588.01	Prostate1 array	2797-2856
DEX0449 047.nt.1	12810	12810.0	Colon array	2773-2832
DEX0449 047.nt.1	31094	31094.02	Prostate2 array	2319-2378
DEX0449 047.nt.2	12810	12810.0	Colon array	2773-2832
DEX0449 047.nt.2	33588	33588.01	Prostate1 array	2797-2856
DEX0449 047.nt.2	31094	31094.02	Prostate2 array	2319-2378
DEX0449 047.nt.2	12809	12809.0	Colon array	2811-2870
DEX0449 048.nt.1	28637	28637.0	Colon array	357-416
DEX0449 048.nt.1	28638	28638.0	Colon array	216-275
DEX0449 049.nt.1	10224	10224.0	Colon array	360-419
DEX0449 049.nt.1	10225	10225.0	Colon array	242-301
DEX0449 051.nt.1	35788	35788.03	Prostate2 array	1107-1166
DEX0449 051.nt.1	28403	28403.0	Colon array	1107-1166
DEX0449 051.nt.1	35746	35746.02	Prostate2 array	893-952
DEX0449 052.nt.1	5820	5820.0	Lung array	1714-1773
DEX0449 052.nt.1	16385	16385.0	Colon array	2216-2275
DEX0449 052.nt.1	5819	5819.0	Lung array	1865-1924
DEX0449 053.nt.1	17655	17655.0	Colon array	1536-1595

DEX0449	053.nt.1	17654	17654.0	Colon array	1600-1659
DEX0449	053.nt.2	17655	17655.0	Colon array	1055-1114
DEX0449	053.nt.2	17654	17654.0	Colon array	1119-1178
DEX0449	054.nt.1	37944	37944.0	Colon array	1296-1355
DEX0449	054.nt.1	6889	6889.0	Lung array	3086-3145
DEX0449	054.nt.1	6890	6890.0	Lung array	2956-3015
DEX0449	054.nt.1	37943	37943.0	Colon array	1416-1475
DEX0449	054.nt.2	6889	6889.0	Lung array	2461-2520
DEX0449	054.nt.2	37944	37944.0	Colon array	671-730
DEX0449	054.nt.2	6890	6890.0	Lung array	2331-2390
DEX0449	055.nt.1	30956	30956.0	Colon array	4398-4457
DEX0449	055.nt.1	31201	31201.0	Colon array	4333-4392
DEX0449	055.nt.1	30951	30951.0	Colon array	4334-4393
DEX0449	055.nt.2	30956	30956.0	Colon array	1809-1868
DEX0449	055.nt.2	31201	31201.0	Colon array	1744-1803
DEX0449	055.nt.3	30951	30951.0	Colon array	3899-3958
DEX0449	055.nt.3	30956	30956.0	Colon array	3963-4022

Example 2b: Relative Quantitation of Gene Expression

Real-Time quantitative PCR with fluorescent Taqman[®] probes is a quantitation detection system utilizing the 5'-3' nuclease activity of Taq DNA polymerase. The method uses an internal fluorescent oligonucleotide probe (Taqman[®]) labeled with a 5' reporter dye and a downstream, 3' quencher dye. During PCR, the 5'-3' nuclease activity of Taq DNA polymerase releases the reporter, whose fluorescence can then be detected by the laser detector of the Model 7700 Sequence Detection System (PE Applied Biosystems, Foster City, CA, USA). Amplification of an endogenous control is used to standardize the amount of sample RNA added to the reaction and normalize for Reverse Transcriptase (RT) efficiency. Either cyclophilin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ATPase, or 18S ribosomal RNA (rRNA) is used as this endogenous control. To calculate relative quantitation between all the samples studied, the target RNA levels for one sample were used as the basis for comparative results (calibrator). Quantitation relative to the "calibrator" can be obtained using the comparative method (User Bulletin #2: ABI PRISM 7700 Sequence Detection System).

The tissue distribution and the level of the target gene are evaluated for every sample in normal and cancer tissues. Total RNA is extracted from normal tissues, cancer tissues, and from cancers and the corresponding matched adjacent tissues. Subsequently, first strand cDNA is prepared with reverse transcriptase and the polymerase chain reaction is done using primers and Taqman[®] probes specific to each target gene. The results are analyzed using the ABI PRISM 7700 Sequence Detector. The absolute numbers are

relative levels of expression of the target gene in a particular tissue compared to the calibrator tissue.

One of ordinary skill can design appropriate primers. The relative levels of expression of the CSNA versus normal tissues and other cancer tissues can then be determined. All the values are compared to the calibrator. Normal RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

The relative levels of expression of the CSNA in pairs of matched samples may also be determined. A matched pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. All the values are compared to the calibrator.

In the analysis of matching samples, the CSNAs show a high degree of tissue specificity for the tissue of interest. These results confirm the tissue specificity results obtained with normal pooled samples. Further, the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual are compared. This comparison provides an indication of specificity for the cancer state (*e.g.* higher levels of mRNA expression in the cancer sample compared to the normal adjacent).

Information on the samples tested in the QPCR experiments include the Sample ID (Smpl ID), Organ, Tissue Type (Tiss Type), Diagnosis (DIAG), Disease Detail, and Stage or Grade (STG or GRD).

Altogether, the high level of tissue specificity, plus the mRNA overexpression in matched samples tested are indicative of SEQ ID NO: 1-112 being a diagnostic marker and/or a therapeutic target for cancer.

Example 3: Protein Expression

The CSNA is amplified by polymerase chain reaction (PCR) and the amplified DNA fragment encoding the CSNA is subcloned in pET-21d for expression in *E. coli*. In addition to the CSNA coding sequence, codons for two amino acids, Met-Ala, flanking the NH₂-terminus of the coding sequence of CSNA, and six histidines, flanking the COOH-terminus of the coding sequence of CSNA, are incorporated to serve as initiating Met/restriction site and purification tag, respectively.

An over-expressed protein band of the appropriate molecular weight may be observed on a Coomassie blue stained polyacrylamide gel. This protein band is confirmed by Western blot analysis using monoclonal antibody against 6X Histidine tag.

Large-scale purification of CSP is achieved using cell paste generated from 6-liter
5 bacterial cultures, and purified using immobilized metal affinity chromatography (IMAC). Soluble fractions that are separated from total cell lysate were incubated with a nickel chelating resin. The column is packed and washed with five column volumes of wash buffer. CSP is eluted stepwise with various concentration imidazole buffers.

Example 4: Fusion Proteins

10 The human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector. For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note
15 that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 2, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced. If the naturally occurring signal sequence is used to
20 produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. *See, e.g.,* WO 96/34891.

Example 5: Production of an Antibody from a Polypeptide

In general, such procedures involve immunizing an animal (preferably a mouse)
25 with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100, µg/ml of streptomycin. The
30 splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP20), available from

the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.*, *Gastroenterology* 80: 225-232 (1981).

The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide. Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

Example 6: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA is isolated from individual patients or from a family of individuals that have a phenotype of interest. cDNA is then generated from these RNA samples using protocols known in the art. *See*, Sambrook (2001), *supra*. The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO: 1-112. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky *et al.*, *Science* 252(5006): 706-9 (1991). *See also* Sidransky *et al.*, *Science* 278(5340): 1054-9 (1997).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons are also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations are then cloned and sequenced to validate the results of the direct sequencing. PCR products is cloned into T-tailed vectors as described in Holton *et al.*, *Nucleic Acids Res.*, 19: 1156

(1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements may also be determined. Genomic clones are nick-translated with digoxigenin deoxyuridine 5' triphosphate (Boehringer Mannheim), and
5 FISH is performed as described in Johnson *et al.*, *Methods Cell Biol.* 35: 73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise mapping
10 are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. Johnson (1991). Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the
15 genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 7: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

Antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably
20 a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described above. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced. The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial
25 dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide. Next, 50 µl of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbound conjugate. 75 µl of
30 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution are added to each well and incubated 1 hour at room temperature.

The reaction is measured by a microtiter plate reader. A standard curve is prepared, using serial dilutions of a control sample, and polypeptide concentrations are plotted on the X-axis (log scale) and fluorescence or absorbance on the Y-axis (linear scale). The concentration of the polypeptide in the sample is calculated using the standard curve.

Example 8: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1, $\mu\text{g/kg/day}$ to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 $\mu\text{g/kg/hour}$ to about 50 mg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U. S. Pat. No.3,773,919, EP 58,481, the contents of which are hereby incorporated by reference herein in their entirety), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22: 547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15: 167-277 (1981), and R. Langer, Chem. Tech. 12: 98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE Epstein et al., Proc. Natl. Acad. Sci. USA 82: 3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77: 4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324, the contents of which are hereby incorporated by reference herein in their entirety. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides. Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably, the carrier is a parenteral carrier, more preferably, a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e. g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1 % (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container (s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 9: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form.

- 5 Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

- For example, a patient with decreased levels of a polypeptide receives a daily dose
10 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided above.

Example 10: Method of Treating Increased Levels of the Polypeptide

- Antisense or RNAi technology are used to inhibit production of a polypeptide of
15 the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

- For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the
20 treatment was well tolerated. The formulation of the antisense polynucleotide is provided above.

Example 11: Method of Treatment Using Gene Therapy

- One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a
25 subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of
30 the flask and fresh media (e. g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 3. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media.

If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 12: Method of Treatment Using Gene Therapy-In Vivo

Another aspect of the present invention is using *in vivo* gene therapy methods to
5 treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide.

The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by
10 the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, Tabata H. *et al. Cardiovasc. Res.* 35 (3): 470-479 (1997); Chao J *et al. Pharmacol. Res.* 35 (6): 517-522 (1997); Wolff J. A. *Neuromuscul. Disord.* 7 (5): 314-318 (1997), Schwartz B. *et al. Gene Ther.* 3 (5): 405-411 (1996); and Tsurumi Y. *et al. Circulation* 94 (12): 3281-3290 (1996); W0 90/11092, W0 98/11779; U. S. Patent No.
15 5,693,622; 5,705,151; 5,580,859, the contents of which are hereby incorporated by reference herein in their entirety.

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, colon, liver, intestine and the like). The polynucleotide
20 constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be
25 delivered in liposome formulations (such as those taught in Felgner P. L. *et al. Ann. NY Acad. Sci.* 772: 126-139 (1995) and Abdallah B. *et al. Biol. Cell* 85 (1): 1-7 (1995)) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain
30 sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the

transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

- The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, colon, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.
- For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 µg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection.
- The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to colons or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle in vivo is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.

The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 13: Transgenic Animals

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e. g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (I. e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection

(Paterson et al., *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver et al., *Biotechnology* 11: 1263-1270 (1993); Wright et al., *Biotechnology* 9: 830-834 (1991); and U. S. Pat. No. 4,873,191, the contents of which is hereby incorporated by reference herein in its entirety); retrovirus mediated gene transfer into germ lines (Van der Putten et al.,
5 *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (Lo, 1983, *Mol Cell. Biol.* 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e. g., Ulmer et al., *Science* 259: 1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells
10 and transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitrano et al., *Cell* 57: 717-723 (1989). For a review of such techniques, see Gordon, "Transgenic Animals," *Intl. Rev. Cytol.* 115: 171-229 (1989).

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated
15 oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., *Nature* 380: 64-66 (1996); Wilmut et al., *Nature* 385: 810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i. e., mosaic animals or chimeric. The transgene may be integrated as a single transgene
20 or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89: 6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest,
25 and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and
30 disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., *Science* 265: 103-106 (1994)). The regulatory sequences required for such a

cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished
5 by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR).
10 Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding
15 strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to
20 both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited
25 to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Example 14: Knock-Out Animals

30 Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E. g., see Smithies et al., Nature 317: 230-234 (1985); Thomas & Capecchi, Cell 51: 503-512 (1987);

Thompson et al., Cell 5: 313-321 (1989)) Alternatively, RNAi technology may be used. For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or
5 without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in
10 research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However, this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to
15 those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e. g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including
20 human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence
25 associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the
30 control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e. g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e. g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U. S. Patent No. 5,399,349; and
5 Mulligan & Wilson, U. S. Patent No. 5,460,959, the contents of which are hereby incorporated by reference herein in their entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the
10 cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of
15 polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

While preferred illustrative embodiments of the present invention are described, one skilled in the art will appreciate that the present invention can be practiced by other
20 than the described embodiments, which are presented for purposes of illustration only and not by way of limitation. The present invention is limited only by the claims that follow.

We claim:

1. An isolated nucleic acid molecule comprising:
 - (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 113-259;
 - 5 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b).
- 10 2. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a cDNA.
- 15 3. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is genomic DNA.
4. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is an RNA.
- 20 5. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a mammalian nucleic acid molecule.
6. The nucleic acid molecule according to claim 5, wherein the nucleic acid molecule is a human nucleic acid molecule.
- 25 7. A method for determining the presence of a colon specific nucleic acid (CSNA) in a sample, comprising the steps of:
 - (a) contacting the sample with the nucleic acid molecule of SEQ ID NO: 1-112 under conditions in which the nucleic acid molecule will selectively hybridize to a colon specific nucleic acid; and
 - 30

(b) detecting hybridization of the nucleic acid molecule to a CSNA in the sample, wherein the detection of the hybridization indicates the presence of a CSNA in the sample.

5 8. A vector comprising the nucleic acid molecule of claim 1.

9. A host cell comprising the vector according to claim 8.

10 10. A method for producing a polypeptide encoded by the nucleic acid molecule according to claim 1, comprising the steps of:

- (a) providing a host cell comprising the nucleic acid molecule operably linked to one or more expression control sequences, and
- (b) incubating the host cell under conditions in which the polypeptide is produced.

15

11. A polypeptide encoded by the nucleic acid molecule according to claim 1.

12. An isolated polypeptide selected from the group consisting of:

- (a) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 113-259 ; or
- (b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112.

20

25 13. An antibody or fragment thereof that specifically binds to:

- (a) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 113-259 ; or
- (b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112.

30

14. A method for determining the presence of a colon specific protein in a sample, comprising the steps of:

- (a) contacting the sample with a suitable reagent under conditions in which the reagent will selectively interact with the colon specific protein comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 113-259; and
- 5 (b) detecting the interaction of the reagent with a colon specific protein in the sample, wherein the detection of binding indicates the presence of a colon specific protein in the sample.
15. A method for diagnosing or monitoring the presence and metastases of colon cancer in a patient, comprising the steps of:
- 10 (a) determining an amount of:
- (i) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 113-259;
- (ii) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112;
- 15 (iii) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (i) or (ii);
- (iv) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (i) or (ii);
- 20 (v) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 113-259 ; or
- (vi) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112
- 25 and;
- (b) comparing the amount of the determined nucleic acid molecule or the polypeptide in the sample of the patient to the amount of the colon specific marker in a normal control; wherein a difference in the amount of the nucleic acid molecule or the polypeptide in the sample compared to the amount of the nucleic acid molecule or the polypeptide in the normal control is associated with the
- 30 presence of colon cancer.

16. A kit for detecting a risk of cancer or presence of cancer in a patient, said kit comprising a means for determining the presence of:
- (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 113-259;
 - 5 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b); or
 - 10 (e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 113-259 ; or
 - (f) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112.
 - 15
17. A method of treating a patient with colon cancer, comprising the step of administering a composition consisting of:
- (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 113-259;
 - 20 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b);
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b);
 - 25 (e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 113-259 ; or
 - (f) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-112;
 - 30
- to a patient in need thereof, wherein said administration induces an immune response against the colon cancer cell expressing the nucleic acid molecule or polypeptide.

18. A vaccine comprising the polypeptide or the nucleic acid encoding the polypeptide of claim 12.

SEQUENCE LISTING

<110> diaDexus, Inc.
Macina, Roberto
Turner, Leah
Sun, Yongming
Rodriguez, Maria

<120> Compositions, Splice Variants and Methods Relating to Colon
Specific Genes and Proteins

<130> DEX-0449

<150> US 60/431,143

<151> 2002-12-04

<150> US 60/431,206

<151> 2002-12-04

<160> 259

<170> PatentIn version 3.1

<210> 1
<211> 1101
<212> DNA
<213> Homo sapien

<400> 1
gggsgccggr rccgcgcccgg gccggctcca gggcgggccgc ggcggtgggg cgcaggccgc 60
cgcgcgagtg aatcgaggcg gcggggcccat ccggaaccgg ccggccatcg cccgoggcgc 120
ggccggcgga ggcggcagga accgaccggc gccctacagc aggccaaaac aacttcccgga 180
caagtggcag cagcatcttt tcgacagtgg ctctggcggt ggtgccggcg tggagacagg 240
tgggaaactg ctggtgtcca atctggattt tggagtctca gacgccgata ttcaggaact 300
ctttgctgaa tttggaacgc tgaagaaggc ggctgtgcac tatgatcgct ctggtcgag 360
cttaggaaca gcagacgtgc actttgagcg gaaggcagat gccctgaagg ccatgaagca 420
gtacaacggc gtccctctgg atgcaagcta catccctccc ctccctcagc tccttccgga 480
agattcactt ctgtgaagcc ttctcagaac cgccatagta cacctctgag tgcggggctg 540
tagtcgtaaa gacaaaggcc gccccatgaa cattcagctt gtcacgtcac agattgacgc 600
acagcggagg cctgcacaga gcgtaaacag aggtggcatg actagaaacc gtggcgctgg 660
aggttttggt ggtggtggag gcacccggag aggcacccgc ggaggcgccc gtggaagagg 720
cagaggtgcc ggcaggaatt caaagcagca gctttcgga gaggagctgg atgccagct 780
ggacgcctat aatgcgagaa tggacaccag ttaaaccagac cagcaaatcc gcgtgcggaa 840
caggaccag gcgtctctc ttgctccctg gttggggggc ggtggctggg gctgtgcggc 900
caatgatgga tttgtttctt ttatgtttta aaataggatt taaaaactca tgtaaagggt 960
tttttttttt cttttttttt ttttttaatt ctgaaacaga cctgttttgt accgagttat 1020

ttttgggata aattttactg gttgctgttg tggagaaggt ggcgtttcca ccttttccat 1080
 aataaaatag aaatgtgtgt a 1101

<210> 2
 <211> 1022
 <212> DNA
 <213> Homo sapien

<400> 2
 cgaacgcctt cgcgcgatcg ccctggaaac gcattctctg cgaccggcag ccgccaatgg 60
 gaagggagtg agtgccacga acaggccaat aaggagggag cagtgcgggg tttaaatctg 120
 aggctaggct ggctcttctc ggcgtgctgc ggcggaacgg ctgttggttt ctgctggttg 180
 taggtccttg gctggtcggg cctcgggtgt tctgcttctc cccgctgagc tgctgcctgg 240
 tgaagaggaa gccatggcgc tccgagtcac caggaaactcg aaaattaatg ctgaaaataa 300
 ggccaagatc aacatggcag gcgcaaagcg cgcttcctacg gcccctgctg caacctccaa 360
 gcccggactg aggccaagaa cagctcttgg ggacattggt aacaaagtca gtgaacaact 420
 gcaggccaaa atgcctatga agaaggaagc aaaaccttca gctactggaa aagtcattga 480
 taaaaaacta ccaaaacctc ttgaaaaggt acctatgctg gtgccagtgc cagtgtctga 540
 gccagtgccg gagccagaac ctgagccaga acctgagcct gttaaagaag aaaaactttc 600
 gcctgagcct attttggttg atactgcctc tccaagccca atggaaacat ctggatgtgc 660
 ccctgcagaa gaagacctgt gtcaggcttt ctctgatgta attcttgagc taaatgatgt 720
 ggatgcagaa gatggagctg atccaaacct ttgtagtga tatgtgaaag atatttatgc 780
 ttatctgaga caacttgagg aagagcaagc agtcagacca aaatacctac tgggtcggga 840
 agtcactgga aacatgagag ccataccta tgaactggtag tacaggttca aatgaaatcc 900
 gtgtgtgcgg gacctgtctg cggtcctatt gacgggcccgg caaaattggg tgcccaaatt 960
 gctggggggc cggccgtttg gcccttgaaa ggccccaggg ggcgttgggg gaccttagca 1020
 ga 1022

<210> 3
 <211> 6046
 <212> DNA
 <213> Homo sapien

<400> 3
 acagtgtcat ggaggccgcg gcggtgacgg tcacccggtc ggctacacgg cggcggcggc 60
 ggcagtgcga ggggctggcg gccccggagg cggggacgca ggaggagcag gaggatcagg 120
 agccgcggcc gcggcggcg cgccgggaa ggagcatcaa agatgaagaa gaagagacag 180

tcttttcgaga ggtgggtcagt ttttccccgg accccctgcc agttagatat tatgacaagg	240
acaccaccaa accaatcagc ttttacttgt cttcgctgga ggagctcttg gcgtggaagc	300
cccgttgga ggatggcttt aatgtggccc tggagccctt ggcgtgtcgc cagccccctc	360
tgagcagcca gaggccccgg actttgttgt gtcatgacat gatgggcggg tacctggatg	420
acaggttcat tcagggtctg gtggtgcaga ctccctatgc tttctaccac tggcagtgca	480
tcgacgtctt tgtgtacttc agccaccaca ccgtcaccat tccccagtg ggctggacca	540
acactgcca caggcatggg gtctgcgtgc tggggacttt catcacggag tggaatgaag	600
ggggaaggct ctgtgaagcc ttcctggccg gggatgagcg ctcgtagcag gcagtggctg	660
accggtggt ccagatcact cagttttttc gttttgatgg ctggctgac aacatcgaga	720
actcgctgag tctggccgct gtggggaaca tgcctccttt cctgcggtac ctcaccacac	780
agctgcaccg gcagggtcca gggggcctgg tgctctggta tgacagcgtg gtgcaaagtg	840
ggcagctcaa atggcaagac gaactcaacc agcacaacag ggtcttcttt gattcctgcg	900
acggcttctt cactaactat aactggcggg aggagcactt ggagcggatg ctggggcagg	960
ctggggagcg cggggtgat gtgtacgtgg gcgtggatgt gtttgctcga gggaacgtgg	1020
tcggaggccg attcgacaca gacaagggtg gtggtggctt tcgtccaagg gccagcggcc	1080
cagtgcctcc cctgggacct catttcctca tggaccttcc attcccgtct gcacctcaaa	1140
ggaatgacag cagctgttct tcccagatg gggatcccgt ggcactgaga aacagggtgc	1200
ctgccccggc caagctatgt cccactgaa acccctttcc cggctatttc tacaatcaga	1260
taatgtattt ttgtgtgggt gttaccagac agaggcgcgt cgttcccccg ctgtccggtc	1320
ctcggcgggc ctgggcctcc tgtcccagct gcgttccttg atcttccccg tgagggaagc	1380
ctgcggggac gcacaggggc ctgcagcgcc tctcactgcg gctcaggggc gaggacgggg	1440
gaggtgcca gggtttgggg atcggtaggt gtcatgacga gggagcattt tctggcaagg	1500
caggtcacac cagcacagag gaaggagcgg ggccaggccg gtgcccctgg cgtgggatgc	1560
gcagggtctt gcctgcctct ggtctatcag tcttgcccgt tcccagagcc ccagggatgt	1620
tctcagccct taatttagct ttaaatgagt aaattaatcc aataactaaa gagcctcttt	1680
actgcccgtg attctgataa tcgaatgaga ggtgtgccag aggacaatgg gaccgcgcag	1740
ccacagccag ttggggccca gccctgccc ccttgccctt gcagcagcgc tgaccgcctc	1800
tcccgcgcc tcttctggcg ggcttgtgga acgacggggg tgatgccag cagcctccag	1860
gtgcatctag gagggagggg ttggacagct tccccctccc gtttctgagc tgccctgaga	1920
gcccgaacct acgtcactgg cgtccaggag cctgtctgga ctctggctct gatgtgcca	1980
gcgtctccct ctgtctcttc agtcgttgga gctgatccga aagcatggct tctccgtggc	2040

tttgtttggc	cccggctggg	tgtatgagtg	tctggagaag	aaggatttct	tccagaacca	2100
ggacaagttc	tggggccgac	tggagcggtta	tctgcccaca	catagcatct	gctccttgcc	2160
tttcgtcacg	tccttctgcc	tgggcatggg	tgcacggagg	gtctgctatg	gccaggaaga	2220
ggcggtaggg	ccctggtacc	acctgagcgc	ccaggagatc	cagcccttgt	ttggagaaca	2280
caggctggga	gggatggcc	ggggctgggt	gaggacgcac	tgctgcctgg	aggatgcctg	2340
gcacggaggc	agctccctgc	tcgtccgggg	tgtgatccca	ccggagggtg	gaaatgtggc	2400
tgtgagggtg	gtgagtgcg	gaggacgggtg	ggcccaccag	cttctcccat	cacacgtggt	2460
ggccatggaa	ctggacagat	ggggcagtg	aggccagaac	aaggacagag	gacagacca	2520
gatgggtttt	ttaaaattgt	gatcaaaaat	gcattaacgt	gaaattcacc	atcttgaccc	2580
tgttgaagtg	aacggtacag	gagtgttaac	tctgcgtacg	ttgctacgtg	acagatccct	2640
agaactgtgt	catcttgtag	gactggactc	tccccattgc	acattagccc	cccatacctc	2700
cccagcccc	caaagccacc	cctttactct	gcactctgtga	gtttggatac	tttagatgcc	2760
ccatatgagg	gaatcccata	cggatatttca	gaactgagat	gtttcagccc	acaaaagaat	2820
ctggccttgg	gggaggactt	gtcccctggc	tacggacagg	accacgatgt	gggagcattt	2880
ggcaccacgg	ccccctgttc	ctggagggag	gggctggtgg	acgcgatttg	gagctggcctt	2940
cgattttctca	gtggtctctc	cacggcaccc	atcacaggac	ctttttcccc	aggttatttt	3000
tattccctgc	aggccccagt	gccacccaag	atttacctgt	ccatggtgta	taagcttgag	3060
gggcccacgg	acgtcacagt	tgctttggag	ctgaccacag	gggatgccgg	cagctgccac	3120
atcgggtggca	tctcagtgtt	gaacgcagaa	acaagctcaa	gacacagcct	ccgacccctc	3180
cgggtgcccc	ccaccaagct	ggccagatgg	gtgggcccgt	gcggccggca	gctgagtggg	3240
ggctgggtcc	agcactgcta	cgagggtgagc	ctgcgtgggt	gcctgctgct	agacctcctc	3300
gtttgcttct	cacggccgcc	gggtagtcgg	gaggaggaga	gcttcacctg	tgggcttggga	3360
gagatccagg	tggtggacgc	tgccagcctg	ctggccccctc	tgccccaggt	gcaggccgtc	3420
accatctctc	acatccgctg	gcagccatcc	gcctctgagc	gggagggggcc	ccctgctctg	3480
ctccagctca	gctgcaccct	gcactgggtcc	ttcctcctct	cacaagtccg	ttgcttcgga	3540
atccactgct	ggggagggat	gagtgatgac	tctccgggca	gggagctgcc	gaggccagag	3600
atgcccattgt	tcctgggggtt	ggcttttgcc	accagtagc	ggatagtggga	cctgctggtg	3660
gaagccgccg	ggcccggcca	ggatcgtcgc	atggaatttc	tggtggagcc	tgtccccaag	3720
gaaggggttcc	gggtacctca	ggccgagtg	ggcagggcag	ttctgcttta	ttcagcccct	3780
gcatagagcg	atgctaaggc	cgggtgggtct	cctggcctcg	ggctgaggcc	tcttcccggc	3840

tgtctgcccc	tggcctgcgc	tggacctgct	aagtgccac	agtggcagcg	aggtcccggg	3900
cccggggctg	gggtgggaga	ccccgggctg	agtgctgtgg	ctttctggtg	gggggagatg	3960
gaaacaggaa	accaagcagt	gggatcgag	cgttggtcac	tgcgaggcga	gtggcgggct	4020
ttctgtttct	gccttgctcc	tccccacggg	acctgggtcc	cagggtgaaa	tgaaaggagg	4080
ggagaagttg	agaacagAAC	attccataaa	ggatatttcc	taataggctg	caagatgctg	4140
atgccgagaa	tgatgatatt	ctttcctgca	gatgaaacta	ttagaaaggg	tcttagattg	4200
tggcaggtag	gctttggagc	aggcgccgag	acatttctga	gcatgaggac	gagctacagc	4260
agctcctggg	gtggggctgc	ctgcgggatg	gcgggagagg	atgccctgga	gaaccgtcct	4320
cccagtgtgg	aaggcccttt	tccctgagga	gtgggcattc	tgggccagcc	ggcgctgggt	4380
tcgctgcctc	acgtgggcca	gccccagctg	ctccgtgttt	cctggcgctg	gcaatttact	4440
gtgctgctga	gtgtgaggtc	atctccggag	cgttttcagc	agccctggc	tctgcccgt	4500
ctcttcggg	ctgtgggcat	gcagggaagt	ggctctgagg	cagtctgcgc	tgtggccctg	4560
cctctgcca	gcgagaggcc	gtgggctctg	gacaagccgc	ccttcaggct	ggggtagcag	4620
gtcagtcag	gcaggaagca	gcacctgccc	cccgccag	cccagcccca	gcctgagtgc	4680
aggagctgca	ggaccgcgg	gggttttcc	agctactctg	ttccttcacg	tcctcccttc	4740
tcagcctcgt	ccaagcaccg	ggaagacctc	caggctgacc	ccttgagcag	cagtcagcac	4800
aggtgcgtgg	ggcggtgagg	gaggcagggt	cttcaccaca	ggcgcttcc	tctgtccttc	4860
ctgctctttc	ttctctgccc	aggccgctgc	agctgcacag	cctctgctac	acctgggctg	4920
cctgggaggc	ttcctggtgt	ggtgtctgga	ccccacggcc	ttgggtcatc	ctgtggctgg	4980
tctgggggtg	ggtctgttgt	ggtccttcca	cgggtgtcagt	ggcctgaagt	ccctcgcttt	5040
tggggggggg	gtctctcacc	cccaggccac	atagggccag	tggtaggggt	tcctctatg	5100
tggggcagtg	ctgagggtg	ggatgctctg	tgacccagc	tggagccac	acctaaagggc	5160
tggcatccac	atcatttcac	cctgcagtga	gggaagaggc	caccagggtg	cagcacagcc	5220
acaccggtc	ccacgtcaga	ggagggcaag	gctgggtact	cagcagccac	tctgagccgg	5280
ggctccttcc	aggagctgaa	atccacctgt	ctccatcttc	cttgctgcc	tgggtactca	5340
tgccaagcag	agactgggat	taggggttct	gtgctcttgc	ctaattagga	acattctccc	5400
atgtctcttg	tgtggtccca	gaaggagaag	tgagtttgcc	aaggatatgg	ggcaggaggc	5460
tcctctgct	gacccctgc	agcctggagc	cagccggggg	actgtcctgg	gtggagggca	5520
ggtgaacaca	agctgctgcc	ggggactgtc	ctgggtggac	ggcaggtgaa	cacaagcggc	5580
tgccgcatgt	agccactcac	tcgacttttt	ttcagctgtg	accattcctg	ggagctcttt	5640
gagcctttct	gtctcatttg	gaaccagggg	gaaccaggaa	ggggctcctg	gcctctctgt	5700

gtcctctgca gtggggggtg tggggggcgc agcatccacg ccttgctgcc cttctttcat 5760
gaagtctgtt ttttaagtgc tggttccccc gaatatatta tgcagaggag ggaaaattta 5820
tagtggcaat tattttctca cagtctggtg agcaggcaat taattaggag taagggggcc 5880
tagtagagcg tggcgtgtgg cagaatcgca ccgccccggc tccccagccc accgccatgc 5940
agggctcgcg tgcgggaaaa ctaatatgcc ggcgtttaag cctgtgcccc tctgctgggt 6000
gtaactgcgc tgaaataaat gatctgacaa tgtgaaaaaa aaaaaa 6046

<210> 4
<211> 5839
<212> DNA
<213> Homo sapien

<400> 4
acagtgtcat ggaggccgcg gcggtgacgg tcacccggtc ggctacacgg cggcggcggc 60
ggcagctgca ggggctggcg gccccggagg cggggacgca ggaggagcag gaggatcagg 120
agccgcggcc gcggcgggcg cggccgggaa ggagcatcaa agatgaagaa gaagagacag 180
tctttcgaga ggtggtcagt ttttccccgg accccctgcc agttagatat tatgacaagg 240
acaccaccaa accaatcagc ttttacttgt cttcgctgga ggagctcttg gcgtggaagc 300
cccgtttgga ggatggcttt aatgtggccc tggagcccct ggcggtgcgc cagccccctc 360
tgagcagcca gaggccccgg actttgttgt gtcatgacat gatgggcggg tacctggatg 420
acaggttcat tcagggtcgc gtggtgcaga ctccctatgc tttctaccac tggcagtgca 480
tcgacgtctt tgtgtacttc agccaccaca ccgtcaccat tccccagtg ggctggacca 540
aactgccc aaggcatggg gtctgcgtgc tggggacttt catcacggag tggaatgaag 600
ggggaaggct ctgtgaagcc ttctggccg gggatgagcg ctctgaccag gcagtggctg 660
accggctggt ccagatcact cagttttttc gttttgatgg ctggctgatc aacatcgaga 720
actcgctgag tctggccgct gtggggaaca tgcctccttt cctgcggtac ctaccacac 780
agctgcaccg gcagggtcca gggggcctgg tgctctggta tgacagcgtg gtgcaaagtg 840
ggcagctcaa atggcaagac gaactcaacc agcacaacag ggtcttcttt gattcctgcg 900
acggcttctt cactaactat aactggcggg aggagcactt ggagcggatg ctggggcagg 960
ctggggagcg ccgggctgat gtgtacgtgg gcgtggatgt gtttgctcga gggaaactgg 1020
tcggaggccg attcgacaca gacaagggtg gtggtggctt tcgtccaagg gccagcggcc 1080
cagtgcctcc cctgggaact catttctca tggaccttcc attcccgtct gcaacctcaa 1140
ggaatgacag cagctgttct tcccagagtg gggatcccg ggactgaga aacagggtgc 1200
ctgccccggc caagctatgt cccactgaa acccctttcc cggctatttc tacaatcaga 1260

taatgtat	ttgtgtgggt	gttaccagac	agaggcgcg	cgttcccccg	ctgtccggtc	1320
ctcggcgggc	ctgggcctcc	tgtcccagct	gcgttccttg	atcttccccg	tgagggaagc	1380
ctgcggggac	gcacaggggc	ctgcagcgcc	tctcactgcg	gctcaggggc	gaggacgggg	1440
gaggtgcca	gggtttgggg	atcggtaggt	gtcatgacga	gggagcattt	tctggcaagg	1500
caggtcacac	cagcacagag	gaaggagcgg	ggccaggccg	gtgcccctgg	cgtgggatgc	1560
gcagggtct	gcctgcctct	ggtctatcag	tcctgcccgt	tcccagagcc	tctttactgc	1620
ccgtgattct	gataatcgaa	tgagaggtgt	gccagaggac	aatgggaccg	cgcagccaca	1680
gccagtggg	gccagcccc	tgccccctg	cccttgacgc	agcgctgacc	gcctctcccc	1740
cgcctcttc	tggcgggctt	gtggaacgac	gggggtgatg	cccagcagcc	tccagtcgtt	1800
ggagctgatc	cgaaagcatg	gcttctccgt	ggctttgttt	gccccgggct	gggtgtatga	1860
gtgtctggag	aagaaggatt	tcttccagaa	ccaggacaag	ttctggggcc	gactggagcg	1920
ttatctgccc	acacatagca	tctgctcctt	gcctttcgtc	acgtccttct	gcctgggcat	1980
gggtgcacgg	agggtctgct	atggccagga	agaggcggt	gggccctgg	accacctgag	2040
cggccaggag	atccagccct	tgtttggaga	acacaggctg	ggaggggatg	gccggggctg	2100
gggtgaggac	cactgctgcc	tggaggatgc	ctggcacgga	ggcagctccc	tgtctgtccg	2160
gggtgtgatc	ccaccggagg	ttggaaatgt	ggctgtgagg	tgggtgagt	acggaggacg	2220
gtggggccac	cagcttctcc	catcacacgt	ggtggccatg	gaactggaca	gatggggcag	2280
tggaggccag	aacaaggaca	gaggacagac	ccagatgggt	tttttaaaat	tgtgatcaaa	2340
aatgcattaa	cgtgaaattc	accatcttga	ccctgttgaa	gtgaacggta	caggagtgtt	2400
aactctgcgt	acgttgctac	gtgacagatc	cctagaactg	tgtcatcttg	taggactgga	2460
ctctcccat	tgcacattag	ccccccatcc	tccccagcc	ccccaaagcc	acccttttac	2520
tctgcatctg	tgagtttgga	tacttttagat	gcccacatg	agggaatccc	atacggtatt	2580
tcagaactga	gatgtttcag	cccacaaaag	aatctggcct	tgggggagga	cttgtcccct	2640
ggctacggac	aggaccacga	tgtgggagca	tttggcaccc	aggccccctg	ttcctggagg	2700
gaggggctgg	tggacgcgat	ttggagctgg	cttcgatttc	tcagtggctc	ctccacggca	2760
cccatcacag	gacctttttc	cccaggttat	ttttattccc	tgcaggcccc	agtgccaccc	2820
aagattttacc	tgtccatgg	gtataagctt	gaggggcccc	cggacgtcac	agttgctttg	2880
gagctgacca	caggggatgc	cggcagctgc	cacatcggtg	gcattctcagt	gttgaacgca	2940
gaaacaagct	caagacacag	cctccgaccc	ctccgggtgc	ccccaccaa	gctggccaga	3000
tgggtgggcc	gctgcggccg	gcagctgagt	gggggctggg	tccagcactg	ctacgaggtg	3060

agcctgcgtg	ggtgcctgct	gctagacctc	ctcgtttget	tctcacggcc	gccgggtagt	3120
cgggaggagg	agagcttcac	ctgtcggctt	ggagagatcc	aggtggtgga	cgctgccagc	3180
ctgctggccc	ctctgcccc	ggtgcaggcc	gtcaccatct	ctcacatccg	ctggcagcca	3240
tccgcctctg	agcgggaggg	gccccctgct	ctgctccagc	tcagctgcac	cctgcactgg	3300
tccttcctcc	tctcacaagt	ccgttgcttc	cgaatccact	gctggggagg	gatgagtgat	3360
gactctccgg	gcaggagct	gccgaggcca	gagatgcccc	tgttcctggg	gttggtttt	3420
gccaccagt	accggatagt	ggacctgctg	gtggaagccg	ccgggcccgg	ccaggatcgt	3480
cgcattggaat	ttctggtgga	gcctgtcccc	aagggaaggt	tccgggtacc	tcaggccgag	3540
tggggcaggg	cagttctgct	ttattcagcc	cctgcatgag	cggatgctaa	ggccgggtgg	3600
tctcctggcc	tcgggctgag	gcctcttccc	ggctgtctgc	ccctggcctg	cgctggacct	3660
gctaagtgcc	cacagtggca	gcgaggtccc	ggtcccgggg	ctgggggtggg	agaccccggg	3720
ctgagtgetg	tggctttctg	gtggggggcg	atggaaacag	gaaaccaagc	agtgggatcg	3780
cagcgttggt	cactgcgagg	cgagtggcgg	gctttctggt	tctgccttgt	ccctccccac	3840
ggtacctggt	tcccaggtga	aaatgaaagg	aggggagaag	ttgagaacag	aacattccat	3900
aaaggatatt	tcctaataag	ctgcaagatg	ctgatgccga	gaatgatgat	tttctttcct	3960
gcagatgaaa	ctattagaaa	gggtcttaga	ttgtggcagg	taggctttgg	agcaggcgcc	4020
gagacatttc	tgagcatgag	gacgagctac	agcagctcct	ggggtggggc	tgcttgcggg	4080
atggcggggg	aggatgccct	ggagaaccgt	cctcccagtg	tgggaaggccc	tttccctga	4140
ggagtgggca	ttctgggcca	gccggcgctg	gcttcgtgcc	tcacagtggg	ccagccccag	4200
ctgctccgtg	tttctggcg	ttggcaattt	actgtgctgc	tgagtgtgag	gtcatctccg	4260
gagcgttttc	agcagcccct	ggctctgcgg	cgtctcttcc	gggctgtggg	catgcaggga	4320
agtggctctg	aggcagtctg	cgctgtggcc	ctgcctctgc	ccagcgagag	gccgtgggct	4380
ctggacaagc	cgcccttcag	gctggggtag	caggtcagtc	caggcaggaa	gcagcacctg	4440
ccccccggcg	cagcccagcc	ccagcctgag	tgcaggagct	gcaggacccg	cgggggcttt	4500
tccagctact	ctgttccttc	acgtcctccc	ttctcagcct	cgtccaagca	ccgggaagac	4560
ctccaggctg	accccttgag	cagcagtcag	cacagggtgcg	tgggggcgtg	agggaggcag	4620
ggtcttcacc	acaggcgctt	tcctctgtcc	ttcctgctct	ttcttctctg	cccaggccgc	4680
tgcagctgca	cagcctctgc	tacacctggg	ctgcctggga	ggcttctctg	tgtggtgtct	4740
ggaccccacg	gccttgggtc	atcctgtggc	tggctctgggg	tggggctctgt	tgtggtcctt	4800
ccacgggtgtc	agtggcctga	agtccctcgc	ttttgggggg	ggggctctctc	accccaggc	4860
cacatagggc	cagtggtagg	ggttccctct	atgtcgggca	gtgctgaggg	ctgggatgct	4920

```

ctgtgacccc agctggagcc cacacctaag ggctggcatc cacatcattt caccctgcag 4980
tgaggggaaga ggccaccagg tggcagcaca gccacacccg ttcccacgtc agaggagggc 5040
aaggctgggt actcagcagc cactctgagc cggggctcct tccaggagct gaaatccacc 5100
tgtctccatc ttcccttgct gcctgggtac tcatgccaaag cagagactgg gattaggggt 5160
tctgtgctct tgcctaatta ggaacattct cccatgtctc ttgtgtggtc ccagaaggag 5220
aagtgagttt gccaaaggata tggggcagga ggctccctct gctgaccccc tgcagcctgg 5280
agccagcccc gggactgtcc tgggtggagg gcaggtgaac acaagctgct gccggggact 5340
gtcctgggtg gacggcaggt gaacacaagc ggctgccga tgtagccact cactcgactt 5400
tttttcagct gtgaccattc ctgggagctc ttgagcctt tctgtctcat ttggaaccag 5460
ggggaaccag gaaggggctc ctggcctctc tgtgtcctct gcagtggggg ttgtgggggg 5520
cgcagcatcc acgccttgct gcccttcttt catgaagtct gttttttaag tgctgggtcc 5580
cccgaatatt ttatgcagag gagggaaaat ttatagtggc aattattttc tcacagtctg 5640
gtgagcaggc aattaattag gagtaagggg gcctagtaga gcgtggcgtg tggcagaatc 5700
gcaccgcccc ggctccccag cccaccgcca tgcagggtc gcgtgcggga aaactaatat 5760
gccggcgttt aagcctgtgc ccctctgtg ggtgtaactg cgctgaaata aatgatctga 5820
caatgtgaaa aaaaaaaaaa 5839

```

```

<210> 5
<211> 5090
<212> DNA
<213> Homo sapien

```

```

<400> 5
acagtgtcat ggaggccgcg gcggtgacgg tcacccggtc ggctacacgg cggcggcggc 60
ggcagctgca ggggctggcg gccccggagg cggggacgca ggaggagcag gaggatcagg 120
agccgcggcc gcggcggcgg cgccgggaa ggagcatcaa agatgaagaa gaagagacag 180
tctttcgaga ggtggtcagt ttttccccgg acccctgcc agttagatat tatgacaagg 240
acaccaccaa accaatcagc ttttacttgt ctctgctgga ggagctcttg gcgtggaagc 300
cccgttgga ggatggcttt aatgtggccc tggagccctt ggctgtctgc cagccccctc 360
tgagcagcca gaggccccgg actttgttgt gtcatgacat gatgggcggg tacctggatg 420
acaggttcat tcagggtctg gtggtgcaga ctccctatgc tttctaccac tggcagtgca 480
tcgacgtctt tgtgtacttc agccaccaca cgtcaccat tccccagtg ggctggacca 540
aactgcccc caggcatggg gtctgcgtgc tggggacttt catcacggag tggaatgaag 600
ggggaaggct ctgtgaagcc ttccctggcg gggtgagcg ctctgaccag gcagtggctg 660

```

accggctggt ccagatcact cagttttttc gttttgatgg ctggctgac aacatcgaga	720
actcgctgag tctggccgct gtggggaaca tgcctccttt cctgcggtac ctcaccacac	780
agctgcaccg gcaggtccca gggggcctgg tgctctggta tgacagcgtg gtgcaaagtg	840
ggcagctcaa atggcaagac gaactcaacc agcacaacag ggtcttcttt gattcctgcg	900
acggcttctt cactaactat aactggcggg aggagcactt ggagcggatg ctggggcagg	960
ctggggagcg ccgggctgat gtgtacgtgg gcgtggatgt gtttgctcga gggaaagtgg	1020
tggaggccg attcgacaca gacaagtcgt tggagctgat ccgaaagcat ggcttctccg	1080
tggctttgtt tgccccggc tgggtgtatg agtgtctgga gaagaaggat ttcttccaga	1140
accaggacaa gttctggggc cgactggagc gttatctgcc cacacatagc atctgctcct	1200
tgcctttcgt cacgtccttc tgcttgggca tgggtgcacg gagggctctg tatggccagg	1260
aagaggcggg agggccctgg taccacctga gcgccaggga gatccagccc ttgtttggag	1320
aacacaggct gggaggggat ggccggggct ggggtaggac gcactgctgc ctggaggatg	1380
cctggcacgg aggcagctcc ctgctcgtcc ggggtgtgat cccaccggag gttggaaatg	1440
tggctgtgag gtgggtgagt gacggaggac ggtgggcca ccagcttctc ccatcacacg	1500
tgggtggccat ggaactggac agatggggca gtggaggcca gaacaaggac agaggacaga	1560
cccagatggg ttttttaaaa ttgtgatcaa aaatgcatta acgtgaaatt caccatcttg	1620
accctgttga agtgaacggg acaggagtgt taactctgcg tacgttgcta cgtgacagat	1680
ccctagaact gtgtcatctt gtaggactgg actctcccca ttgcacatta gcccccatc	1740
ctccccagc cccccaaagc cacccttta ctctgcatct gtgagtttgg atactttaga	1800
tgccccatat gagggaatcc catacggtat ttcagaactg agatgtttca gccacaaaa	1860
gaatctggcc ttgggggagg acttgctccc tggctacgga caggaccacg atgtgggagc	1920
atttggcacc caggccccct gttcctggag ggaggggctg gtggacgcga tttggagctg	1980
gcttcgattt ctcagtggc tctccacggc acccatcaca ggaccttttt ccccaggtta	2040
tttttatccc ctgcaggccc cagtgccacc caagatttac ctgtccatgg tgtataagct	2100
tgaggggccc acggacgtca cagttgcttt ggagctgacc acaggggatg ccggcagctg	2160
ccacatcggg ggcattctcag tgttgaacgc agaaacaagc tcaagacaca gcctccgacc	2220
cctccgggtg cccccacca agctggccag atgggtgggc cgctgcggcc ggcagctgag	2280
tgggggctgg gtccagcact gctacgaggt gagcctgcgt gggcgctgc tgctagacct	2340
cctcgtttgc ttctcacggc cgccgggtag tcgggaggag gagagcttca cctgtcggct	2400
tggagagatc caggtggtgg acgctgccag cctgctggcc cctctgcccc aggtgcaggc	2460

11

cgtcaccatc tctcacatcc gctggcagcc atccgcctct gagcgggagg ggccccctgc	2520
tctgctccag ctgagctgca ccctgcactg gtcccttctc ctctcacaag tccgttgctt	2580
ccgaatccac tgctggggag ggatgagtga tgactctccg ggcagggagc tgccgaggcc	2640
agagatgccc atgttcctgg ggttggtctt tgccaccagc taccggatag tggacctgct	2700
ggtggaagcc gccggggccg gccaggatcg tcgcatggaa tttctggtgg agcctgtccc	2760
caaggaaggg ttccgggtac ctgaggccga gtggggcagg gcagttctgc tttattcagc	2820
ccctgcatga gcggatgcta aggccgggtg gtctcctggc ctggggctga ggccctctcc	2880
cggctgtctg cccctggcct gcgctggacc tgctaagtgc ccacagtggc agcgaggtcc	2940
cggctccggg gctgggggtg gagaccccg gctgagtgtc gtggctttct ggtggggggc	3000
gatggaaaca ggaaaccaag cagtgggatc gcagcgttg tccactgcag gcgagtggcg	3060
ggctttctgt ttctgccttg tccctcccca cggtagcttg ttcccagggtg aaaatgaaag	3120
gaggggagaa gttgagaaca gaacattcca taaaggatat ttccataatg gctgcaagat	3180
gctgatgccg agaatgatga ttttctttcc tgcagatgaa actattagaa agggctcttag	3240
attgtggcag gtaggctttg gagcaggcgc cgagacattt ctgagcatga ggacgagcta	3300
cagcagctcc tgggggtggg ctgcctgcgg gatggcggga gaggatgccc tggagaaccg	3360
tcctcccagt gtggaaggcc cttttccctg aggagtgggc attctgggccc agccggcgct	3420
ggcttcgtgc ctccacgtgg gccagcccca gctgctccgt gtttcctggc gttggcaatt	3480
tactgtgctg ctgagtgtga ggtcatctcc ggagcgtttt cagcagcccc tggctctgcg	3540
gcgtctcttc cgggctgtgg gcatgcaggg aagtggctct gaggcagtct gcgctgtggc	3600
cctgcctctg cccagcgaga ggccgtgggc tctggacaag ccgcccttca ggctggggta	3660
gcaggtcagt ccaggcagga agcagcacct gcccccgcg ccagcccagc cccagcctga	3720
gtgcaggagc tgcaggaccc gcgggggctt ttccagctac tctgttcctt cactgcctcc	3780
cttctcagcc tcgtccaagc accgggaaga cctccaggct gacctcttga gcagcagtca	3840
gcacagggtg gtgggggcgt gagggaggca gggctctcac cacaggcgcc ttctctgtc	3900
cttctgtctc tttcttctct gccaggccg ctgcagctgc acagcctctg ctacacctgg	3960
gctgcctggg aggccttctg gtgtggtgtc tggacccac gcccttgggt catcctgtgg	4020
ctggctctgg gtgggtctct ttgtggtcct tccacggtgt cagtggcctg aagtccctcg	4080
cttttggggg gggggtctct cccccagg ccacataggg ccagtggtag gggttccctc	4140
tatgtcgggc agtgtgagg gctgggatgc tctgtgacct cagctggagc ccacacctaa	4200
gggctggcat ccacatcatt tcacctgca gtgagggaag aggccaccag gtggcagcac	4260
agccacaccc gttcccacgt cagaggaggg caaggctggg tactcagcag ccactctgag	4320

```

ccgggggctcc ttccaggagc tgaaatccac ctgtctccat cttecttgcc tgcctgggta 4380
ctcatgccaa gcagagactg ggattagggg ttctgtgctc ttgcctaatt aggaacattc 4440
tcccatgtct cttgtgtggt ccagaaagga gaagtgagtt tgccaaggat atggggcagg 4500
aggtccctc tgctgacccc ctgcagcctg gagccagccc ggggactgtc ctgggtggag 4560
ggcaggtgaa cacaagctgc tgccggggac tgcctgggt ggacggcagg tgaacacaag 4620
cggtgccgc atgtagccac tactcgact ttttttcagc tgtgaccatt cctgggagct 4680
ctttgagcct ttctgtctca tttggaacca gggggaacca ggaaggggct cctggcctct 4740
ctgtgtcctc tgcagtgggg gttgtgggg ggcagcatc cacgccttg tgccttctt 4800
tcatgaagtc tgttttttaa gtgctgggtc cccgaatat tttatgcaga ggagggaaaa 4860
tttatagtgg caattatctt ctcacagtct ggtgagcagg caattaatta ggagtaaggg 4920
ggcctagtag agcgtggcgt gtggcagaat cgcaccgcc cggtcccca gccaccgcc 4980
atgcagggct cgcgtgcggg aaaactaata tgccggcgtt taagcctgtg cccctctgct 5040
gggtgtaact gcgctgaaat aaatgatctg acaatgtgaa aaaaaaaaaa 5090

```

```

<210> 6
<211> 5621
<212> DNA
<213> Homo sapien

```

```

<400> 6
gcggcgggtgg cgggaaggag catcaaagat gaagaagaag agacagtctt tcgagaggtg 60
gtcagttttt ccccggaacc cctgccagtt agatattatg acaaggacac caccaaacca 120
atcagctttt acttgtcttc gctggaggag ctcttggcgt ggaagccccg cttggaggat 180
ggctttaatg tggccctgga gcccctggcg tgcgccagc cccctctgag cagccagagg 240
ccccggactt tgttgtgtca tgacatgatg ggcggtacc tggatgacag ggactttcat 300
cacggagtgg aatgaagggg gaaggctctg tgaagccttc ctggccgggg atgagcgctc 360
gtaccaggca gtggtgacc ggctggcca gatcactcag ttttttcgtt ttgatggctg 420
gctgatcaac atcgagaact cgctgagtct ggccgctgtg gggaacatgc ctctttcct 480
gcggtacctc accacacagc tgcaccggca ggtcccagg ggcctggtgc tctggtatga 540
cagcgtggtg caaagtgggc agctcaaatg gcaagacgaa ctcaaccagc acaacagggt 600
cttctttgat tcctgcgacg gcttcttcac taactataac tggcgggagg agcacttgga 660
gcggatgctg gggcaggctg gggagcgccg ggctgatgtg tacgtgggcg tggatgtgtt 720
tgctcgaggg aacgtggtcg gaggccgatt cgacacagac aaggtgggtg gtggctttcg 780
tccaagggcc agcggcccag tgcctcccct gggacctcat ttcctcatgg acctccatt 840

```

cccgtctgca	cctcaaagga	atgacagcag	ctgttcttcc	cagagtgggg	atcccgtggc	900
actgagaaac	aggtgtcctg	ccccggccaa	gctatgtccc	cactgaaacc	cctttcccgg	960
ctatttctac	aatcagataa	tgtatttttg	tgtgggtggt	accagacaga	ggcgcgtcgt	1020
tcccccgctg	tccggtcctc	ggcgggcctg	ggcctcctgt	cccagctgcg	ttccttgatc	1080
ttccccgtga	gggaagcctg	cggggacgca	caggggcctg	cagcgcctct	cactgaggct	1140
caggggcgag	gacgggggag	gtgcccagg	tttggggatc	ggtaggtgtc	atgacgaggg	1200
agcattttct	ggcaaggcag	gtcacaccag	cacagaggaa	ggagcggggc	caggccgggtg	1260
cccctggcgt	gggatgcgca	gggctctgcc	tgcctctggt	ctatcagtcc	tgcccgttcc	1320
cagagcccca	gggatgttct	cagcccttaa	tttagcttta	aatgagtaaa	ttaatccaat	1380
aactaaagag	cctctttact	gcccgtgatt	ctgataatcg	aatgagaggt	gtgccagagg	1440
acaatgggac	cgcgcagcca	cagccagttg	gggccagacc	cctgcccccc	tgcccttgca	1500
gcagcgctga	ccgcctctcc	cgcgccctct	tctggcgggc	ttgtggaacg	acgggggtga	1560
tgcccagcag	cctccagtcg	ttggagctga	tccgaaagca	tggcttctcc	gtggctttgt	1620
ttgcccccg	ctgggtgtat	gagtgtctgg	agaagaagga	tttcttccag	aaccaggaca	1680
agttctgggg	ccgactggag	cgttatctgc	ccacacatag	catctgctcc	ttgcctttcg	1740
tcacgtcctt	ctgcctgggc	atgggtgcac	ggagggtctg	ctatggccag	gaagaggcgg	1800
tagggccctg	gtaccacctg	agcgcccagg	agatccagcc	cttgtttgga	gaacacaggc	1860
tgggagggga	tggccggggc	tgggtgagga	cgcactgctg	cctggaggat	gcctggcacg	1920
gaggcagctc	cctgctcgtc	cggggtgtga	tcccaccgga	ggttggaat	gtggctgtga	1980
ggtgggtgag	tgacggagga	cgggtggccc	accagcttct	cccatcacac	gtggtggcca	2040
tggaaactgga	cagatggggc	agtggaggcc	agaacaagga	cagaggacag	accagatgg	2100
gttttttaaa	attgtgatca	aaaatgcatt	aacgtgaaat	tcaccatctt	gaccctgttg	2160
aagtgaacgg	tacaggagtg	ttaactctgc	gtacgttgct	acgtgacaga	tccctagaac	2220
tgtgtcatct	tgtaggactg	gactctcccc	attgcacatt	agccccccat	cctccccag	2280
ccccccaaag	ccaccctttt	actctgcata	tgtgagtttg	gatactttag	atgccccata	2340
tgagggaatc	ccatacggta	tttcagaact	gagatgtttc	agcccacaaa	agaatctggc	2400
cttgggggag	gacttgtccc	ctggctacgg	acaggaccac	gatgtgggag	catttggcac	2460
ccaggccccc	tgttcctgga	gggaggggct	ggtggacgcg	atttggagct	ggcttcgatt	2520
tctcagtgg	ctctccacgg	caccatcac	aggacctttt	tccccagggt	atttttatte	2580
cctgcaggcc	ccagtgccac	ccaagattta	cctgtccatg	gtgtataagc	ttgaggggcc	2640

cacggacgtc acagttgctt tggagctgac cacaggggat gccggcagct gccacatcgg	2700
tggcatctca gtgttgaacg cagaaacaag ctcaagacac agcctccgac ccctccgggt	2760
gccccccacc aagctggcca gatgggtggg ccgctgcggc cggcagctga gtgggggctg	2820
ggtccagcac tgctacgagg tgagcctgcg tgggtgcctg ctgctagacc tcctcgtttg	2880
cttctcacgg ccgccgggta gtcgggagga ggagagcttc acctgtcggc ttggagagat	2940
ccagggtggtg gacgtgcca gcctgctggc ccctctgccc cagggtgcagg ccgtcaccat	3000
ctctcacatc cgctggcagc catccgcctc tgagcgggag gggccccctg ctctgctcca	3060
gctcagctgc acctgcact ggtccttcct cctctcacia gtccgttgct tccgaatcca	3120
ctgctgggga gggatgagt atgactctcc gggcagggag ctgccgaggc cagagatgcc	3180
catgttcctg gggttggctt ttgccacca gtaccggata gtggacctgc tggtggaagc	3240
cgccggggcc ggccaggatc gtcgcatgga atttctggtg gagcctgtcc ccaaggaagg	3300
gttccgggta cctcaggccg agtggggcag ggcagttctg ctttattcag cccctgcatg	3360
agcggatgct aaggccgggt ggtctcctgg cctcgggctg aggcctcttc ccggctgtct	3420
gccccctggc tgcgctggac ctgctaagtg cccacagtgg cagcgaggtc ccggtcccgg	3480
ggctggggtg ggagaccccc ggctgagtgc tgtggctttc tggtgggggg cgatggaaac	3540
aggaacacaa gcagtgggat cgcagcgttg gtactgcca ggcgagtggc gggctttctg	3600
tttctgcctt gtccctcccc acggtacctg gttcccaggt gaaaatgaaa ggaggggaga	3660
agttgagaac agaacattcc ataaaggata tttcctaata ggctgcaaga tgctgatgcc	3720
gagaatgatg attttctttc ctgcagatga aactattaga aagggtctta gattgtggca	3780
ggtaggcttt ggagcaggcg ccgagacatt tctgagcatg aggacgagct acagcagctc	3840
ctgggggtggg gctgcctgcg ggatggcggg agaggatgcc ctggagaacc gtcctcccag	3900
tgtggaaggc ccttttccct gaggagtggg cattctgggc cagccggcgc tggcttcgtg	3960
cctccacgtg ggccagcccc agctgctccg tgtttcctgg cgttggcaat ttactgtgct	4020
gctgagtgtg aggtcatctc cggagcgttt tcagcagccc ctggctctgc ggcgtctctt	4080
ccgggctgtg ggcattgcagg gaagtggctc tgaggcagtc tgcgctgtgg ccctgcctct	4140
gcccagcgag aggccgtggg ctctggacaa gccgcccttc aggctggggg agcaggtcag	4200
tccaggcagg aagcagcacc tgccccccgc gccagcccag cccagcctg agtgcaggag	4260
ctgcaggacc cgcgggggct tttccageta ctctgttctc tcacgtcctc ccttctcagc	4320
ctcgtccaag caccgggaag acctccaggc tgaccccttg agcagcagtc agcacaggtg	4380
cgtggggggc tgaggagggc agggctctca ccacaggcgc cttcctctgt ccttctctgt	4440
ctttcttctc tgcccaggcc gctgcagctg cacagcctct gctacacctg ggctgcctgg	4500

15

```

gaggcttcct ggtgtggtgt ctggacccca cggccttggg tcatcctgtg gctgggtctgg 4560
ggtgggggtct gttgtggtcc ttccacgggtg tcagtggcct gaagtccctc gcttttgggg 4620
gggggggtctc tcacccccag gccacatagg gccagtggta ggggttccct ctatgtcggg 4680
cagtgtctgag ggctgggatg ctctgtgacc ccagctggag cccacaccta agggctggca 4740
tccacatcat ttaccctgc agtgaggga gaggccacca ggtggcagca cagccacacc 4800
cgttcccacg tcagaggagg gcaaggctgg gtactcagca gccactctga gccggggctc 4860
cttccaggag ctgaaatcca cctgtotcca tcttccctgc ctgcctgggt actcatgcca 4920
agcagagact gggattagg gttctgtgct cttgcctaatt taggaacatt ctcccatgtc 4980
tcttgtgtgg tcccagaagg agaagtgagt ttgccaagga tatggggcag gaggctccct 5040
ctgctgaccc cctgcagcct ggagccagcc cggggactgt cctgggtgga gggcaggtga 5100
acacaagctg ctgccgggga ctgtcctggg tggacggcag gtgaacacaa gcggctgccg 5160
catgtagcca ctactcgac tttttttcag ctgtgaccat tcctgggagc tctttgagcc 5220
tttctgtctc atttgaacc agggggaacc aggaaggggc tcctggcctc tctgtgtcct 5280
ctgcagtggg ggttgtgggg ggcgcagcat ccacgccttg ctgcccttct ttcattgaagt 5340
ctgtttttta agtgctggtt cccccaata ttttatgcag aggagggaaa atttatagtg 5400
gcaattatct tctcacagtc tggtagcag gcaattaatt aggagtaagg gggcctagta 5460
gagcgtggcg tgtggcagaa tcgcaccgcc cgggctcccc agcccaccgc catgcagggc 5520
tcgcgtgagg gaaaactaat atgccggcgt ttaagcctgt gccctctgc tgggtgtaac 5580
tgcgctgaaa taaatgatct gacaatgtga aaaaaaaaaa a 5621

```

```

<210> 7
<211> 5519
<212> DNA
<213> Homo sapien

```

```

<400> 7
acagtgtcat ggaggccgcg gcggtgacgg tcacccggtc ggctacacgg cggcggcggc 60
ggcagctgca ggggctggcg gccccggagg cggggacgca ggaggagcag gaggatcagg 120
agccgcggcc gcggcggcgg cggccgggaa ggagcatcaa agatgaagaa gaagagacag 180
tctttcgaga ggtggtcagt ttttccccgg accccctgcc agttagatat tatgacaagg 240
acaccaccaa accaatcagc ttttacttgt ctctcgctgga ggagctcttg gcgtggaagc 300
cccgtttgga ggatggcttt aatgtggccc tggagccctt ggcgtgtcgc cagccccctc 360
tgagcagcca gaggccccgg actttgttgt gtcatgacat gatgggcggg tacctggatg 420
acaggttcat tcagggctcg gtggtgcaga ctccctatgc tttctaccac tggcagtgca 480

```

tcgacgtctt	tgtgtacttc	agccaccaca	ccgtcaccat	tccccagtg	ggctggacca	540
acactgcca	caggcatggg	gtctgcgtgc	tggggacttt	catcacggag	tggaatgaag	600
ggggaaggct	ctgtgaagcc	ttcctggccg	gggatgagcg	ctcgtaccag	gcagtggctg	660
accggctggg	ccagatcact	cagttttttc	gttttgatgg	ctggctgatc	aacatcgaga	720
actcgctgag	tctggccgct	gtggggaaca	tgcctccttt	cctgcggtac	ctcaccacac	780
agctgcaccg	gcagggtccca	gggggcctgg	tgctctggta	tgacagcgtg	gtgcaaagtg	840
ggcagctcaa	atggcaagac	gaactcaacc	agcacaacag	ggtcttcttt	gattcctgcg	900
acggcttctt	cactaactat	aactggcggg	aggagcactt	ggagcggatg	ctggggcagg	960
ctggggagcg	ccgggctgat	gtgtacgtgg	gcgtggatgt	gtttgctcga	gggaacgtgg	1020
tcggaggccg	attcgacaca	gacaagggtg	gtgggtggctt	tcgtccaagg	gccagcggcc	1080
cagtgcctcc	cctgggacct	catttcctca	tggaccttcc	attcccgtct	gcacctcaaa	1140
ggaatgacag	cagctgttct	tcccagagtg	gggatcccgt	ggcactgaga	aacagggtgtc	1200
ctgccccggc	caagctatgt	ccccactgaa	acccctttcc	cggctatttc	tacaatcaga	1260
taatgtat	ttgtgtgggt	gttaccagac	agaggcgct	cgttcccccg	ctgtccggtc	1320
ctcggcgggc	ctgggcctcc	tgtcccagct	gcgttccttg	atcttccccg	tgagggaagc	1380
ctgcggggac	gcacaggggc	ctgcagcgcc	tctcactgcg	gctcaggggc	gaggacgggg	1440
gagggtgcca	gggtttgggg	atcggtaggt	gtcatgacga	gggagcattt	tctggcaagg	1500
cagggtcacac	cagcacagag	gaaggagcgg	ggccaggccg	gtgcccctgg	cgtgggatgc	1560
gcagggtctt	gcctgcctct	ggtctatcag	tcctgcccgt	tcccagagcc	ccagggtatgt	1620
tctcagccct	taatttagct	ttaaagtagt	aaattaatcc	aataactaaa	gagcctcttt	1680
actgcccgtg	attctgataa	tcgaatgaga	ggtgtgccag	aggacaatgg	gaccgcgcag	1740
ccacagccag	ttggggccca	gcccctgccc	ccctgcccct	gcagcagcgc	tgaccgcctc	1800
tcccgcgcc	tcttctggcg	ggcttgtgga	acgacggggg	tgatgccag	cagcctccag	1860
tcgttggagc	tgatccgaaa	gcatggcttc	tccgtggctt	tgtttgcccc	cggctgggtg	1920
tatgagtgtc	tggagaagaa	ggatttcttc	cagaaccagg	acaagtctctg	gggccgactg	1980
gagcgttatc	tgcccacaca	tagcatctgc	tccttgccct	tcgtcacgtc	cttctgcctg	2040
ggcatgggtg	cacggagggt	ctgctatggc	caggaagagg	cggtagggcc	ctgggtaccac	2100
ctgagcgccc	aggagatcca	gcccttgttt	ggagaacaca	ggctgggagg	ggatggccgg	2160
ggctgggtga	ggacgcactg	ctgcctggag	gatgcctggc	acggaggcag	ctccctgctc	2220
gtccgggggtg	tgatcccacc	ggagggttga	aatgtggctg	tgagggtgggt	gagtgcgga	2280

17

ggacggtggg	cccaccagct	tctcccatca	cacgtggtgg	ccatggaact	ggacagatgg	2340
ggcagtggag	gccagaacaa	ggacagagga	cagaccaga	tgggtttttt	aaaattgtga	2400
tcaaaaatgc	attaacgtga	aattcaccat	cttgaccctg	ttgaagtga	cggtagagga	2460
gtgttaactc	tgcgtacgtt	gctacgtgac	agatccctag	aactgtgtca	tcttgttagga	2520
ctggactctc	ccattgcac	attagcccc	catcctcccc	cagcccccca	aagccacccc	2580
tttactctgc	atctgtgagt	ttggatactt	tagatgcccc	atatgaggga	atccatacag	2640
gtatttcaga	actgagatgt	ttcagcccac	aaaagaatct	ggccttgggg	gaggacttgt	2700
cccctggcta	cggacaggac	cacgatgtgg	gagcatttgg	caccagggcc	ccctgttctc	2760
ggagggaggg	gctggtggac	gcgatttggg	gctggcttcg	atttctcagt	ggtctctcca	2820
cggcacccat	cacaggacct	ttttccccag	gttattttta	ttccctgcag	gccccagtgc	2880
caccaagat	ttacctgtcc	atggtgtata	agcttgaggg	gcccacggac	gtcacagtgt	2940
ctttggagct	gaccacaggg	gatgccggca	gctgccacat	cggtaggcac	tcagtgttga	3000
acgcagaaac	aagctcaaga	cacagcctcc	gacccctccg	ggtgcccccc	accaagctgg	3060
ccagatgggt	gggccgctgc	ggccggcagc	tgagtggggg	ctgggtccag	cactgctacg	3120
aggtgagcct	gcgtgggtgc	ctgctgctag	acctcctcgt	ttgcttctca	cggccgcagg	3180
gtagtcggga	ggaggagagc	ttcacctgtc	ggcttgagg	gatccaggtg	gtggacgctg	3240
ccagcctgct	ggccccctctg	ccccaggtgc	aggccgtcac	catctctcac	atccgctggc	3300
agccatccgc	ctctgagcgg	gaggggcccc	ctgctctgct	ccagctcagc	tgcacctgc	3360
actggtcctt	cctcctctca	caagtccgtt	gcttccgaat	ccactgctgg	ggagggatga	3420
gtgatgactc	tccgggcagg	gagctgccga	ggccagagat	gcccattgtc	ctgggggttg	3480
cttttgccac	ccagtaccgg	atagtggacc	tgctggtgga	agccgcccgg	cccggccagg	3540
atcgctcgcat	ggaatttctg	gtggagcctg	tccccaagga	agggttccgg	gtacctcagg	3600
ccgagtgggg	cagggcagtt	ctgctttatt	cagcccctgc	atgagcggat	gctaaggccg	3660
ggtggtctcc	tggcctcggg	ctgaggcctc	ttcccggctg	tctgcccctg	gcctgcgctg	3720
gacctgctaa	gtgccacag	tggcagcgag	gtcccggctc	cggggctggg	gtgggagacc	3780
ccgggctgag	tgctgtggct	ttctgggtgg	gggagatgga	aacaggaaaac	caagcagtgg	3840
gatcgacgag	ttggtcactg	cgaggcgagt	ggcgggcttt	ctgtttctgc	cttgtccctc	3900
cccacggtac	ctggttccca	ggtgaaaatg	aaaggagggg	agaagttag	aacagaacat	3960
tccataaagg	atatttccta	ataggctgca	agatgctgat	gccgagaatg	atgattttct	4020
ttcctgcaga	tgaactatt	agaaagggtc	ttagattgtg	gcaggtaggc	tttgagcag	4080
gcgccgagac	atttctgagc	atgaggacga	gctacagcag	ctcctggggg	ggggctgcct	4140

gcgggatggc gggagaggat gccctggaga accgtcctcc cagtgtggaa ggcccttttc 4200
 cctgaggagt gggcattctg ggccagccgg cgctggcttc gtgcctccac gtgggccagc 4260
 cccagctgct ccgtgtttcc tggcgttggc aatttactgt gctgctgagt gtgaggcat 4320
 ctccggagcg ttttcagcag cccctggctc tgcggcgtct cttccgggct gtgggcatgc 4380
 aggggaagtgg ctctgaggca gtctgcgctg tggccctgcc tctgcccagc gagaggccgt 4440
 gggctctgga caagccgccc ttcaggctgg ggtagcaggt cagtccaggc aggaagcagc 4500
 acctgcccc cgcgccagcc cagccccagc ctgagtgcag gagctgcagg acccgcgggg 4560
 gcttttccag ctactctgtt ccttcacgtc ctcccttctc agcctcgtcc aagcaccggg 4620
 aagacctcca ggctgacccc ttgagcagca gtcagcacag tgagggaaga ggccaccagg 4680
 tggcagcaca gccacacccg tttccacgtc agaggagggc aaggctgggt actcagcagc 4740
 cactctgagc cggggctcct tccaggagct gaaatccacc tgtctccatc ttccttgctc 4800
 gcctgggtac tcatgccaag cagagactgg gattaggggt tctgtgctct tgcctaatta 4860
 ggaacattct cccatgtctc ttgtgtggtc ccagaaggag aagtgagttt gccaaaggata 4920
 tggggcagga ggctccctct gctgacccc tgcagcctgg agccagcccg gggactgtcc 4980
 tgggtggagg gcagggtgaac acaagctgct gccggggact gtcctgggtg gacggcaggt 5040
 gaacacaagc ggctgccgca tgtagccact cactcgactt tttttcagct gtgaccattc 5100
 ctgggagctc tttgagcctt tctgtctcat ttggaaccag ggggaaccag gaaggggctc 5160
 ctggcctctc tgtgtcctct gcagtggggg ttgtgggggg cgagcatcc acgccttgct 5220
 gcccttcttt catgaagtct gttttttaag tgctggttcc ccgaatatt ttatgcagag 5280
 gagggaaaat ttatagtggc aattattttc tcacagtctg gtgagcaggc aattaattag 5340
 gagtaagggg gcctagtaga gcgtggcgtg tggcagaatc gcaccgcccc ggctccccag 5400
 cccaccgcca tgcagggtc gcgtgcggga aaactaatat gccggcggtt aagcctgtgc 5460
 ccctctgctg ggtgtaactg cgctgaaata aatgatctga caatgtgaaa aaaaaaaaa 5519

<210> 8

<211> 6102

<212> DNA

<213> Homo sapien

<400> 8

acagtgtcat ggaggccgcg gcggtgacgg tcaccgggtc ggctacacgg cggcgggcggc 60
 ggcagctgca ggggctggcg gccccggagg cggggacgca ggaggagcag gaggatcagg 120
 agccgcggcc gcgscggcgg cggccgggaa ggagcatcaa agatgaagaa gaagagacag 180
 tctttcagaa ggtggtcagt ttttccccgg accccctgcc agttagatat tatgacaagg 240

acaccaccaa	accaatcagc	ttttacttgt	cttcgctgga	ggagctcttg	gcgtggaagc	300
cccgtcttga	ggatggcttt	aatgtggccc	tggagccctt	ggcgtgtcgc	cagccccctc	360
tgagcagcca	gaggccccgg	actttgttgt	gtcatgacat	gatgggcggg	tacctggatg	420
acaggttcat	tcagggtctg	gtggtgcaga	ctccctatgc	tttctaccac	tggcagtgca	480
tcgacgtctt	tgtgtacttc	agccaccaca	cgtcaccat	tccccagtg	ggctggacca	540
acactgcca	caggcatggg	gtctgcgtgc	tggggacttt	catcacggag	tggaatgaag	600
ggggaaggct	ctgtgaagcc	ttcctggccg	gggatgagcg	ctcgtaccag	gcagtggctg	660
accggctggg	ccagatcact	cagttttttc	gttttgatgg	ctggctgata	aacatcgaga	720
actcgtgag	tctggccgct	gtggggaaca	tgcctccttt	cctgcggtac	ctcaccacac	780
agctgcaccg	gcagggtcca	gggggcctgg	tgctctggta	tgacagcgtg	gtgcaaagtg	840
ggcagctcaa	atggcaagac	gaactcaacc	agcacaacag	ggtcttcttt	gattcctgcg	900
acggcttctt	cactaactat	aactggcggg	aggagcactt	ggagcggatg	ctggggcagg	960
ctggggagcg	ccgggctgat	gtgtacgtgg	gcgtggatgt	gtttgctcga	gggaacgtgg	1020
tcggaggccg	attcgacaca	gacaagggtg	gtggtggctt	tcgtccaagg	gccagcggcc	1080
cagtgcctcc	cctgggacct	catttcctca	tggaccttcc	attcccgtct	gcacctcaaa	1140
ggaatgacag	cagctgttct	tcccagagtg	gggatcccgt	ggcactgaga	aacagggtgc	1200
ctgccccggc	caagctatgt	ccccactgaa	acccctttcc	cggctatttc	tacaatcaga	1260
taatgtatth	ttgtgtgggt	gttaccagac	agaggcgctg	cgttcccccg	ctgtccggctc	1320
ctcggcgggc	ctgggcctcc	tgtcccagct	gcgttccttg	atcttccccg	tgagggaagc	1380
ctgcggggac	gcacaggggc	ctgcagcgcc	tctcactgcg	gctcaggggc	gaggacgggg	1440
gaggtgccca	gggtttgggg	atcggtaggt	gtcatgacga	gggagcattt	tctggcaagg	1500
caggtcacac	cagcacagag	gaaggagcgg	ggccaggccg	gtgcccctgg	cgtgggatgc	1560
gcagggtctt	gcctgcctct	ggtctatcag	tcctgcccgt	tcccagagcc	ccagggatgt	1620
tctcagccct	taatttagct	ttaaatagag	aaattaatcc	aataactaaa	gagcctcttt	1680
actgcccgtg	attctgataa	tcgaatgaga	ggtgtgccag	aggacaatgg	gaccgcgcag	1740
ccacagccag	ttggggccca	gcccccgccc	ccctgccttt	gcagcagcgc	tgaccgcctc	1800
tcccgcgccc	tcttctggcg	ggcttgtgga	acgacggggg	tgatgccccag	cagcctccag	1860
tcgttggagc	tgatccgaaa	gcatggcttc	tccgtggctt	tgtttgcccc	cggtggggtg	1920
tatgagtgtc	tggagaagaa	ggatttcttc	cagaaccagg	acaagtcttg	gggccgactg	1980
gagcgttatc	tgcccacaca	tagcatctgc	tccttgccct	tcgtcacgtc	cttctgcctg	2040

20

ggcatgggtg	cacggagggg	ctgctatggc	caggaagagg	cggtagggcc	ctggtaccac	2100
ctgagcgccc	aggagatcca	gcccttggtt	ggagaacaca	ggctgggagg	ggatggccgg	2160
ggctgggtga	ggacgcactg	ctgcctggag	gatgcctggc	acggaggcag	ctccctgctc	2220
gtccgggggtg	tgatcccacc	ggaggttgga	aatgtggctg	tgagggtggg	gagtgcagga	2280
ggacgggtggg	cccaccagct	tctcccatca	cacgtgggtg	ccatggaact	ggacagatgg	2340
ggcagtggag	gccagaacaa	ggacagagga	cagaccacga	tgggtttttt	aaaattgtga	2400
tcaaaaatgc	attaacgtga	aattcaccat	cttgaccctg	ttgaagtga	cggtagagga	2460
gtgttaactc	tgcgtacgtt	gctacgtgac	agatccctag	aactgtgtca	tcttgtagga	2520
ctggactctc	cccattgcac	attagcccc	catectcccc	cagcccccca	aagccacccc	2580
tttactctgc	atctgtgagt	ttggatactt	tagatgcccc	atatgaggga	atcccatacg	2640
gtatttcaga	actgagatgt	ttcagcccac	aaaagaatct	ggccttgggg	gaggacttgt	2700
cccctggcta	cggacaggac	cacgatgtgg	gagcatttgg	caccagggcc	ccctgttctc	2760
ggagggaggg	gctggtggac	gcgatttggg	gctggcttcg	atttctcagt	ggtctctcca	2820
cggcacccat	cacaggacct	ttttccccag	gttattttta	ttccctgcag	gccccagtgc	2880
cacccaagat	ttacctgtcc	atggtgtata	agcttgaggg	gccacaggac	gtcacagttg	2940
ctttggagct	gaccacaggg	gatgccggca	gctgccacat	cggtaggcac	tcagtgttga	3000
acgcagaaac	aagctcaaga	cacagcctcc	gaaccctccg	ggtgcccccc	accaagctgg	3060
ccagatgggt	gggccgctgc	ggccggcagc	tgagtggggg	ctgggtccag	caactgctacg	3120
aggtgagcct	gcgtgggtgc	ctgctgctag	acctcctcgt	ttgcttctca	cggccgcccgg	3180
gtagtgggga	ggaggagagc	ttcacctgtc	ggcttggaga	gatccagggt	atgcttccca	3240
gaggggctcg	ggctgggctg	gctgtttgtc	cagctggagt	gggggtggag	gccgccccag	3300
gccgccccct	tcttgggttc	agcggggaac	tggggtggag	gagtcaggga	ggggaaatgt	3360
gtgcgtgggg	gcacctctc	cccgcccccg	ggcgtccagc	cgtgctgagc	cttctctcct	3420
gccaggtggt	ggacgctgcc	agcctgctgg	cccctctgcc	ccaggtgcag	gccgtcacca	3480
tctctcacat	cgcgtggcag	ccatccgcct	ctgagcggga	ggggccccct	gctctgctcc	3540
agctcagctg	caccctgcac	tggtccttcc	tctctcaca	agtccgttgc	ttccgaatcc	3600
actgctgggg	agggatgagt	gatgactctc	cgggcaggga	gctgccgagg	ccagagatgc	3660
ccatgttctc	ggggttggct	tttgccaccc	agtaccggat	agtggacctg	ctggtggaag	3720
ccgccggggc	cggccaggat	cgtcgcctgg	aatttctggg	ggagcctgtc	cccaagggaag	3780
ggttccgggt	acctcaggcc	gagtggggca	gggcagttct	gctttattca	gcccctgcat	3840
gagcggatgc	taaggccggg	tggctctcctg	gcctcgggct	gaggcctctt	cccggctgtc	3900

tgccctggc ctgcgctgga cctgctaagt gccacagtg gcagcgaggt cccggtccc	3960
gggctggggg gggagacccc gggctgagtg ctgtggcttt ctgggtgggg gcgatggaaa	4020
caggaaacca agcagtggga tcgcagcgtt ggtcactgcg aggcgagtg cgggctttct	4080
gtttctgcct tgtccctccc cacggtacct ggttcccagg tgaaaatgaa aggaggggag	4140
aagttgagaa cagaacattc cataaaggat atttcctaag aggctgcaag atgctgatgc	4200
cgagaatgat gattttcttt cctgcagatg aaactattag aaagggctct agattgtggc	4260
aggtaggctt tggagcaggc gccgagacat ttctgagcat gaggacgagc tacagcagct	4320
cctgggggtg ggctgcctgc gggatggcgg gagaggatgc cctggagaa cgtcctccca	4380
gtgtggaagg cccttttccc tgaggagtgg gcattctggg ccagccggcg ctggcttcgt	4440
gcctccacgt gggccagccc cagctgctcc gtgtttcctg gcgttgga tttactgtgc	4500
tgctgagtg gaggtcatct cggagcgtt ttcagcagcc cctggctctg cggcgtctct	4560
tccgggctgt gggcatgcag ggaagtggct ctgaggcagt ctgcgctgtg gccctgcctc	4620
tgccagcga gaggcgtgg gctctggaca agccgccctt caggctgggg tagcaggcca	4680
gtccaggcag gaagcagcac ctgccccccg ccgagccca gcccagcct gaggcagga	4740
gctgcaggac ccgcgggggc ttttccagct actctgttcc ttcacgtcct cccttctcag	4800
cctcgtccaa gcaccgggaa gacctccagg ctgaccctt gagcagcagt cagcacagg	4860
gcgtgggggc gtgagggagg cagggtcttc accacaggcg ccttctctg tcttctctg	4920
tctttctct ctgcccaggc cgtgcagct gcacagcctc tgctacacct gggctgcctg	4980
ggaggcttcc tgggtgtgtg tctggacccc acggccttgg gtcactcctg ggctggtctg	5040
gggtggggtc tgttgtgtgc cttccacggt gtcagtggcc tgaagtccct cgcttttggg	5100
gggggggtct ctacccccca ggccacatag ggccagtggg aggggttccc tctatgtcgg	5160
gcagtgctga gggctgggat gctctgtgac ccagctgga gcccacacct aagggtggc	5220
atccacatca tttcacctg cagtgagga agaggccacc aggtggcagc acagccacac	5280
ccgttccac gtcagaggag ggcaaggctg ggtactcagc agccactctg agccggggct	5340
ccttccagga gctgaaatcc acctgtctcc atcttccttg cctgcctggg tactcatgcc	5400
aagcagagac tgggattagg ggttctgtgc tcttgctaa ttaggaacat tctcccatgt	5460
ctcttgtgtg gtcccagaag gagaagttag tttgccaagg atatggggca ggaggctccc	5520
tctgctgacc ccctgcagcc tggagccagc ccggggactg tcctgggtgg agggcagggtg	5580
aacacaagct gctgccgggg actgtcctgg gtggacggca ggtgaacaca agcggctgcc	5640
gcatgtagcc actcactcga ctttttttca gctgtgacca ttcctgggag ctctttgagc	5700

22

ctttctgtct catttggaac cagggggaac caggaagggg ctctggcct ctctgtgtcc 5760
 tctgcagtgg gggttgtggg gggcgagca tccacgcctt gctgcccttc tttcatgaag 5820
 tctgtttttt aagtgtggt tccccgaat attttatgca gaggagggaa aatttatagt 5880
 ggcaattatt ttctcacagt ctggtgagca ggcaattaat taggagtaag ggggcctagt 5940
 agagcgtggc gtgtggcaga atcgaccgc cccggctccc cagcccaccg ccatgcaggg 6000
 ctgcgctgcg ggaaaactaa tatgccggcg tttaagcctg tgcccctctg ctgggtgtaa 6060
 ctgcgctgaa ataatgatc tgacaatgtg aaaaaaaaaa aa 6102

<210> 9
 <211> 3177
 <212> DNA
 <213> Homo sapien

<400> 9
 acagtgtcat ggaggccgcg gcggtgacgg tcaccgggtc ggctacacgg cggcggcggc 60
 ggcagctgca ggggctggcg gccccggagg cggggacgca ggaggagcag gaggatcagg 120
 agccgcggcc gggcgggcg cgccgggaa ggagcatcaa agatgaagaa gaagagacag 180
 tctttcgaga ggtggtcagt ttttccccgg accccctgcc agttagatat tatgacaagg 240
 acaccaccaa accaatcagc ttttacttgt cttegtgga ggagctcttg gcgtggaagc 300
 cccgcttgga ggatggcttt aatgtggccc tggagccctt ggctgtctgc cagccccctc 360
 tgagcagcca gaggccccgg actttgttgt gtcatgacat gatgggcggg tacctggatg 420
 acaggttcat tcagggtctg gtggtgcaga ctccctatgc tttctaccac tggcagtgca 480
 tcgacgtctt tgtgtacttc agccaccaca ccgtcaccat tccccagtg ggctggacca 540
 aactgcccc caggcatggg gtctgcgtgc tggggacttt catcacggag tggaatgaag 600
 ggggaaggct ctgtgaagcc ttctggccg gggatgagcg ctctaccag gcagtggctg 660
 accggtggt ccagatcact cagtttttct gttttgatgg ctggctgate aacatcgaga 720
 actcgtgag tctggccgct gtggggaaca tgcctccttt cctgcggtac ctaccacac 780
 agctgctggt ggaagccgcc gggccggcc aggatcgtcg catggaattt ctggtggagc 840
 ctgtcccaa ggaagggttc cgggtacctc aggccgagt gggcagggca gttctgcttt 900
 attcagcccc tgcagagcg gatgctaagg ccgggtggtc tcctggcctc gggctgaggc 960
 ctcttccccg ctgtctgccc ctggcctgcg ctggacctgc taagtgccca cagtggcagc 1020
 gaggtcccgg tcccggggct ggggtgggag accccgggct gagtgtgtg gctttctggt 1080
 ggggggagat ggaaacagga aaccaagcag tgggatcgca gcgttggtca ctgcgaggcg 1140
 agtggcgggc tttctgttct tgccttgctc ctcccacgg tacctggttc ccaggtgaaa 1200

atgaaaggag gggagaagtt gagaacagaa cattccataa aggatatttc ctaataggct	1260
gcaagatgct gatgccgaga atgatgattt tctttcctgc agatgaaact attagaaagg	1320
gtcttagatt gtggcaggta ggctttggag caggcgccga gacatttctg agcatgagga	1380
cgagctacag cagctcctgg ggtggggctg cctgcgggat ggcgggagag gatgccctgg	1440
agaaccgtcc tcccagtgtg gaaggccctt ttccctgagg agtgggcatt ctggggccagc	1500
cggcgtggc ttctgtcctc cacgtgggcc agccccagct gctccgtgtt tcctggcggt	1560
ggcaatttac tgtgtgctg agtgtgaggt catctccgga gcgttttcag cagccccctg	1620
ctctgcgccg tctcttcctg gctgtgggca tgcagggaag tggctctgag gcagtctgcg	1680
ctgtggccct gcctctgccc agcgagaggc cgtgggctct ggacaagccg cccttcaggc	1740
tggggtagca ggtcagtcca ggagggaagc agcacctgcc ccccgcgcca gccagcccc	1800
agcctgagtg caggagctgc aggaccgcg ggggcttttc cagctactct gttccttcac	1860
gtcctccctt ctcagcctcg tccaagcacc gggaagacct ccaggctgac cccttgagca	1920
gcagtcagca cagggtgcgtg ggggcgtgag ggaggcaggg tcttcaccac aggcgccttc	1980
ctctgtcctt cctgtctctt cttctctgcc caggccgtg cagctgcaca gcctctgcta	2040
cacctgggct gcctgggagg cttcctgggt tgggtgtctg accccacggc cttgggtcat	2100
cctgtggctg gtctggggtg gggctctgtt tggctccttc acgggtgtcag tggcctgaag	2160
tccctcgctt ttgggggggg ggtctctcac ccccaggcca catagggcca gtggtagggg	2220
ttccctctat gtggggcagt gctgagggtt gggatgctct gtgacccag ctggagccca	2280
cacctaaggg ctggcatcca catcatttca ccctgcagtg aggggaagagg ccaccagggtg	2340
gcagcacagc cacaccggtt cccacgtcag aggagggcaa ggctgggtac tcagcagcca	2400
ctctgagccg gggctccttc caggagctga aatccacctg tctccatctt ccttgccctgc	2460
ctgggtactc atgccaagca gagactggga ttaggggttc tgtgtctctg cctaattagg	2520
aacattctcc catgtctctt gtgtgggtccc agaaggagaa gtgagtttgc caaggatatg	2580
gggcaggagg ctccctctgc tgacccctg cagcctggag ccagcccggg gactgtcctg	2640
ggtggagggc aggtgaacac aagctgctgc cggggactgt cctgggtgga cggcagggtga	2700
acacaagcgg ctgcccatg tagccaactc ctcgactttt ttccagctgt gaccattcct	2760
gggagctctt tgagcctttc tgtctcattt ggaaccaggg ggaaccagga aggggctcct	2820
ggcctctctg tgcctctgc agtgggggtt gtggggggcg cagcatccac gccttgctgc	2880
ccttctttca tgaagtctgt tttttaagtg ctgggtcccc cgaatatatt atgcagagga	2940
gggaaaattt atagtggcaa ttattttctc acagtctggt gagcaggcaa ttaattagga	3000
gtaagggggc ctagtagagc gtggcgtgtg gcagaatcgc accgccccgg ctccccagcc	3060

caccgccatg cagggctcgc gtgcgggaaa actaatatgc cggcgtttaa gcctgtgccc 3120
 ctctgctggg tgtaactgcg ctgaaataaa tgatctgaca atgtgaaaaa aaaaaaa 3177

<210> 10
 <211> 1956
 <212> DNA
 <213> Homo sapien

<400> 10
 ccggctccac ccccaagcca ggcgaggcag gttccgaggt tggaacacct ggcgagtcct 60
 cgggtgctcggg ggccggcagc catctcgcgg ccgttcagga taagccagag acatgggaaa 120
 accaatggaa gtgtttcaaa gttcttctct ttaatcactt ctgtatccag ctgcctttga 180
 tttgtggaac ctattatatt acagagtatt tcaatattcc ttatgattgg gaaagaatgc 240
 caagatggta ttttcttttg gcaagatgct ttggttgtgc agtcattgaa gatacttggc 300
 actattttct gcatagactc ttacaccaca aaagaatata caagtatatt cataaagttc 360
 atcatgagtt tcaggctcca tttggaatgg aagctgaata tgcacatcct ttggagactc 420
 taattcttgg aactggattt ttcatggaa tcgtgctttt gtgtgatcat gtaattcttc 480
 tttgggcatg ggtgaccatt cgtttattag aaactattga tgtccatagt ggttatgata 540
 ttcctctcaa ccctttaaat ctgatccctt tctatgctgg ttctcggcat catgatttcc 600
 accacatgaa cttcattgga aactatgctt caacatttac atgggtgggat cgaatttttg 660
 gaacagactc tcagtataat gcctataatg aaaagaggaa gaagtttgag aaaaagactg 720
 aataaatatc tcacgtaaac cttcctgaaa gataaacgtt ttctgaatt cagaaactag 780
 tagctaacat tgcttctgga gagcagaaat aagcatgtct tctggctact aagtataaa 840
 aagaacatta acaaccttta attaccttc tagtgggaac tttttctact ttacctaca 900
 gttctatata ttagaaaatg aataaatata tatttaagta cagttttcat gaggaagttt 960
 taaaagacca tgttcctaag cttccaagaa ggttttggat actagaagta ttaatctatg 1020
 gcttttctcc cagtaaaacc ataggcctga agttcacatt gggctcttaa atctttttaga 1080
 tatatactgg tcatttcaga aaattcttca tagtgggtatt ggccttatat ttaacttttt 1140
 ttttattttt tttttgagac aaagccacac tctgtctcct tggctggagt gtggtggcac 1200
 agtctcagct cactgcaacc tctgcctccc agttcaagca attcttctgc ctcagcctcc 1260
 caagtagctg ggattacagg caccgccac caccgccagc taatttttgt atttttgtag 1320
 agatgggggt tcacgatgtt ggccaggctg gtctcaaact tctgacctca agtgatctgc 1380
 ccaccttggc ctcccaaagt gctgggatta caggtgtaag ccactgcgcc cggccttttt 1440
 aactttaaac atgttttaga attcacctaa agatcaaaat atcatggatt gaacctcatc 1500

25

```

aattgatagc agtgagtgac tgaagcttcc aaatcaagaa aagccggcac caagaacttc 1560
cattctaatac tagagctgac cagtttgagc tgattctctc tttgaagagt ccttcttgat 1620
tgacgtgcag tactggcatt tctgaatgga tgtaagtgga gtatttttagt ctaaaggctt 1680
ttcaaattac ttgaattttt ttaaaaattg aggagcttta tttctattta cccttccatt 1740
tttgtatatac aaatttccat tgtcattaaa aactgtatct tgaaactttg tgaactgact 1800
tgctgtatct gcactttgag ctcttgaaat aaatgtgatt tttgtgtgat tatctggttt 1860
ccagttttta acattaactg tcacctttta ttcttaaact tgaaagtaca gaaatcatta 1920
aattattaag ttgtacaata aagggtattg tttctt 1956

```

```

<210> 11
<211> 1488
<212> DNA
<213> Homo sapien

```

```

<400> 11
gactccaagt cccagctgtg cccgggacgc tgcgcgcccc cgggactcca tttcccagag 60
tgccccgccc cagcctcccg gccccgcccc ccacctggct ccgcgggcag ccggagcgca 120
cccggccgga agccgctgtc ggggagccgg cgggtggggct ggacgcagggt gcaactgaca 180
tggggtgaacc ccagggatcc atgcggattc tagtgacagg gggctctggg ctggtaggca 240
aagccatcca gaagggtgga gcagatggag ctggacttcc tggagaggac tgggtgtttg 300
tctcctctaa agacgccgat ctacgcgata cagcacagac ccgcgccctg tttgagaagg 360
tccaacccac acacgtcatc catcttgtct caatgggtggg gggcctgttc cggaatatca 420
aatacaatct ggacttctgg aggaaaaacg tgcacatgaa cgacaacgtc ctgcactcgg 480
ccttcgagggt gggcgcccg c aagggtgtgt cctgcctgtc cacctgtatc ttccctgaca 540
agacgacctc cccgatagat gagaccatga tccacaatgg gcctccccac aacagcaatt 600
ttgggtactc gtatgccaaag aggatgatcg acgtgcagaa cagggcctac ttccagcagt 660
acggctgcac ctacaccgct gtcacccca ccaacgtctt cgggccccac gacaacttca 720
acatcgagga tggccacgtg ctgcctggcc tcatccaaa ggtgcacctg gccaaagca 780
gcggctcggc cctgacgggtg tggggtacag ggaatccgcg gaggcagttc atatactcg 840
tggaacctggc ccagctcttt atctgggtcc tgcgggagta caatgaagtg gagcccatca 900
tcctctccgt gggcgaggaa gatgaggtct ccatcaagga ggcagccgag gcgggtgttg 960
aggccatgga ctcccatggg gaagtcacct ttgatacaac caagtcggat gggcagttta 1020
agaagacagc cagtaacagc aagctgagga cctacctgcc cgacttccgg ttcacaccct 1080
tcaagcaggc ggtgaaggag acctgtgctt ggttactga caactacgag caggcccgga 1140

```

26

```

agtgaagctg gaagacagga tcaggtgcca gcggaccatc ggctggcaga gccagcggc 1200
caccacccgt caaccctgcc aggagctgag ggcaccaccc agcaacctgg gcctgcattc 1260
catccgctct gcagcccca gcatctttcc agtggggccc ccattcacgt tggctctcag 1320
ggaaaccagg gtccggggca gggccggcgc tttgctcccc acaccagccc cctgcgcgtg 1380
tccactctga tcctgcatcc cactccctgg gagccaataa agtgcatttt cacaggcaaa 1440
aaaggggaag aaaacgacag ggggaggtag aaggagtagg aggagtga 1488

```

```

<210> 12
<211> 1757
<212> DNA
<213> Homo sapien

```

```

<400> 12
gactccaagt cccagctgtg cccgggacgc tgcgcgcccc cgggactcca tttcccagag 60
tgccccgccc cagcctcccg gccccgcccc ccacctggct ccgcgggcag ccggagcgca 120
cccggccgga agccgctgtc ggggagccgg cgggtggggct ggacgcaggt gcaactgaca 180
tgggtgaacc ccagggatcc atgcggattc tagtgacagg gggctctggg ctggtaggca 240
aagccatcca gaaggtggta gcagatggag ctggacttcc tggagaggac tgggtgtttg 300
tctcctctaa agacgccgat ctcacggata cagcacagac ccgcgccctg tttgagaagg 360
tccaaccac acacgtcatc catcttgtct caatgggtgg gggcctgttc cggaatatca 420
aatacaatth ggacttctgg aggaaaaacg tgcacatgaa cgacaacgtc ctgcactcgg 480
ccttcgaggt gggcgcccg c aaggtggtgt cctgcctgtc cacctgtatc ttccctgaca 540
agacgacctt cccgatagat gagaccatga tccacaatgg gcctccccac aacagcaatt 600
ttgggtactc gtatgccaa aggatgatcg acgtgcagaa cagggcctac ttccagcagt 660
acggctgcac cttcaccgct gtcacccca ccaacgtctt cgggccccac gacaacttca 720
acatcgagga tggccacgtg ctgcctggcc tcatccacaa ggtgcacctg gccaagagca 780
gcggctcggc cctgacgggtg tggggtagag ggaatccgcg gaggcagttc atatactcg 840
tggaacctgg ccagctcttt atctgggtcc tgcgggagta caatgaagtg gagcccatca 900
tcctctccgt gggcgaggaa gatgaggtct ccatcaagga ggcagccgag gcgggtggtg 960
aggccatgga cttccatggg gaagtcacct ttgatacaac caagtcggat gggcagttta 1020
agaagacagc cagtaacagc aagctgagga cctacctgcc cgacttccgg ttcacaccct 1080
tcaagcaggc ggtgaaggag acctgtgctt ggttcactga caactacgag caggcccgga 1140
agtgaagctg gaagacagga tcaggtgcca gcggaccatc ggctggcaga gccagcggc 1200
caccacccgt caaccctgcc aggagctgag ggcaccaccc agcaacctgg gcctgcattc 1260

```

27

catccgctct	gcagcccaaa	gcatctttcc	agtggggccc	ccattcacgt	tggtcctcag	1320
ggaaaccagg	gtccggggca	ggcccggcgc	tttgetcccc	acaccagccc	cctgcgcgtg	1380
tccactctga	tcctgcatcc	cactccctgg	gagccaataa	agtgcatttt	cacaggcaaa	1440
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaagggg	gggggcgcca	1500
aaagattccc	cctcgggggg	ggccgccaca	atattatgcc	gcagcagccc	acggtgtttc	1560
tgtgtgtgaa	aaagaagggg	gggcccgaat	aaaggggggc	gcgtaatata	tgaacccccg	1620
acgagatggg	cgcccggctt	taaaaaaacg	cggcgggtga	gtgggcgaaa	aaacgcgtga	1680
taccggggca	aacctgggga	gaagagacac	ccccaacac	ggaggggggg	agaccaatcg	1740
ggggcaacac	actacct					1757

<210> 13
 <211> 1869
 <212> DNA
 <213> Homo sapien

<400> 13	
gggagccgcg	cgacgcctgc
gcagggcccc	ctcccctaag
ggctggccct	ggggccctgc
60	
cgggggagcc	cgagcaggtt
tccggctccc	aggatgggat
ccttgagggt	ggctgcgccc
120	
cgaggccctgt	cctgccctgg
aggaatctgg	ctcgccaggg
ctctccagga	ccagccagcc
180	
tcgagtgggc	gtgggctgca
cacaggcgcc	gtctcgtgcc
tgtccctagc	atgcgggctt
240	
tagctgcgaa	ccgtgttaac
gacctgtgcc	aggagccccc
gtcccaagg	tgtctgcctc
300	
ccctcttgtc	agtcagagg
gagtagaatg	taccttttct
cggcctagt	gggagtcgtg
360	
ggttgggacc	agctgctctg
ggctgggtgg	tagctcagga
cccctgcggc	gctgtaggct
420	
gagagctccc	aggggaacag
ggctgcgacg	agggagtgcc
tcggtccagc	tgggtctttc
480	
tggttgccag	tggactatgc
cccactctga	gggcctcact
ctgtgccagt	tgccacagaa
540	
gagtggggct	ccaaaggatg
agagcgtgat	gacctcagcc
tcccactgtt	tgacattaca
600	
gtgcaactga	catgggtgaa
ccccagggat	ccatgcggat
tctagtgaca	gggggctctg
660	
ggctggtagg	caaagccatc
cagaagggtg	tagcagatgg
agctggactt	cctggagagg
720	
actgggtggt	tgtctcctct
aaagacgccg	atctcacgga
tacagcacag	acccgcgccc
780	
tgtttgagaa	ggtccaaccc
acacacgtca	tccatcttgc
tgcaatggtg	gggggcctgt
840	
tccggaatat	caaatacaat
ttggacttct	ggaggaaaaa
cgtgcacatg	aacgacaacg
900	
tcctgcactc	ggccttcgag
gtgggcgccc	gcaagggtgg
gtcctgcctg	tccacctgta
960	
tcttcctga	caagacgacc
tacccgatag	atgagacat
gatccacaat	gggcctcccc
1020	
acaacagcaa	ttttgggtac
tcgtatgcca	agaggatgat
cgacgtgcag	aacagggcct
1080	

```

acttccagca gtacggctgc accttcaccc ctgtcatccc caccaacgtc ttcgggcccc 1140
acgacaactt caacatcgag gatggccacg tgctgcctgg cctcatccac aagggtgcacc 1200
tggccaagag cagcggtctg gccctgacgg tgtgggggtac agggaatccg cggaggcagt 1260
tcatatactc gctggacctg gccagctct ttatctgggt cctgcgggag tacaatgaag 1320
tggagcccat catcctctcc gtgggcgagg aagatgaggt ctccatcaag gaggcagccg 1380
aggcgggtgt ggaggccatg gacttccatg ggaagtcac ctttgatata accaagtcgg 1440
atgggcagtt taagaagaca gccagtaaca gcaagctgag gacctacctg cccgacttcc 1500
ggttcacacc cttcaagcag gcggtgaagg agacctgtgc ttggttcact gacaactacg 1560
agcaggcccc gaagtgaagc tggaagacag gatcaggtgc cagcggacca tcggctggca 1620
gagcccagcg gccaccaccc gtcaacctg ccaggagctg agggcaccac ccagcaacct 1680
gggcctgcat tccatccgct ctgcagcccc aagcatcttt ccagtggggc cccattcac 1740
gttggtcttc agggaaacca ggggtccggg caggccccggc gctttgctcc ccacaccagc 1800
cccctgcgcg tgtccactct gatcctgcat cccactccct gggagccaat aaagtgcatt 1860
ttcacaggc 1869

```

```

<210> 14
<211> 1602
<212> DNA
<213> Homo sapien

```

```

<400> 14
gactccaagt cccagctgtg cccgggacgc tgcgcgcccc cgggactcca tttcccagag 60
tgccccgccc cagcctcccg gccccgcccc ccacctggct ccgcgggcag ccggagcgca 120
cccggccgga agccgctgtc ggggagccgg cgggtggggct ggacgcaggt gcaactgaca 180
tgggtgaacc ccagggatcc atgcggattc tagtgacagg gggctctggg ctggtaggca 240
aagccatcca gaaggtggta gcagatggag ctggacttcc tggagaggac tgggtgtttg 300
tctcctctaa agacgccgat ctcacggata cagcacagac ccgcgccctg tttgagaagg 360
tccaaccac acacgtcatc catcttgctg caatgggtggg gggcctgttc cggaatatca 420
aatacaattt ggacttctgg aggaaaaacg tgcacatgaa cgacaacgtc ctgcactcgg 480
ccttcgaggt gggcgccccg aagggtgggt cctgcctgtc cacctgtatc ttccctgaca 540
agaagaccta cccgatagat gagaccatga tccacaatgg gcctccccac aacagcaatt 600
ttgggtactc gtatgccaag aggatgatcg acgtgcagaa caggctctcg cggccatcct 660
gctgaggccc agccaggagt cggggagagg gccccaggg gcgggtgggg ccggctgggc 720
gggtcttcgt gccagcctca gtaagaggcc agcggcagcc tggggtggga tcgcggggccc 780

```

```

aggacgaggg caggagacgc tgtgcctcgg cctacttcca gcagtacggc tgcaccttca 840
ccgctgtcat cccacccaac gtcttcgggc ccacgacaa cttcaacatc gaggatggcc 900
acgtgctgcc tggcctcatc cacaaggtgc acctggccaa gagcagcggc tcggccctga 960
cgggtgtgggg tacaggggaat ccgcggaggg agttcatata ctcgctggac ctggcccagc 1020
tctttatctg ggtcctgcgg gagtacaatg aagtggagcc catcatcctc tccgtgggcg 1080
aggaagatga ggtctccatc aaggaggcag ccgaggcggg ggtggaggcc atggacttcc 1140
atggggaagt cacctttgat acaaccaagt cggatgggca gtttaagaag acagccagta 1200
acagcaagct gaggacctac ctgcccgaact tccggttcac acccttcaag caggcgggtga 1260
aggagacctg tgcttggttc actgacaact acgagcaggg ccggaagtga agctggaaga 1320
caggatcagg tgccagcggg ccatcggctg gcagagccca ggggccacca cccgtcaacc 1380
ctgccaggag ctgagggcac caccagcaa cctgggcctg cattccatcc gctctgcagc 1440
cccaagcatc ttccagtggt ggccccatt cacgttggtc ctcagggaaa ccagggtccg 1500
gggcaggccc ggcgctttgc tccccacacc agccccctgc gcgtgtccac tctgatcctg 1560
catcccactc cctgggagcc aataaagtgc attttcacag gc 1602

```

```

<210> 15
<211> 1693
<212> DNA
<213> Homo sapien

```

```

<400> 15
gactccaagt cccagctgtg cccgggacgc tgcgcgcccc cgggactcca tttcccagag 60
tgccccgccc cagcctcccg gccccgcccc ccacctggct ccgcgggcag ccggagcgca 120
cccggccgga agccgctgtc ggggagccgg cgggtggggct ggacgcaggt gcaactgaca 180
tgggtgaacc ccagggatcc atgcggattc tagtgacagg gggctctggg ctggtaggca 240
aagccatcca gaaggtggta gcagatggag ctggacttcc tggagaggac tgggtgtttg 300
tctcctctaa agacgccgat ctacggata cagcacagac ccgcgccctg tttgagaagg 360
tccaaccac acacgtcatc catcttgctg caatggtggg gggcctgttc cggaatatca 420
aatacaattt ggacttctgg aggaaaaacg tgcacatgaa cgacaacgtc ctgcactcgg 480
ccttcgaggt gggcgccccg aagggtggtg cctgcctgtc cacctgtatc ttccctgaca 540
agacgacctc cccgatagat gagacatga tccacaatgg gcctccccac aacagcaatt 600
ttgggtactc gtatgccaa aggatgatcg acgtgcagaa caggtcctcg cgcccatcct 660
gctgaggccc agccaggagt cggggagagg gccccaggg gcgggtgggg ccggtctggc 720
gggtcttcgt gccagcctca gtaagaggcc agcggcagcc tggggtggga tcgcgggccc 780

```

```

aggacgaggg caggagacgc tgtgcctcgg cctacttcca gcagtacggc tgcaccttca      840
ccgctgtcat cccaccaaac gtcttcgggc cccacgacaa cttcaacatc gaggatggcc      900
acgtgctgcc tggcctcatc cacaaggtgc acctggccaa gagtgagtgc gtccgtccga      960
tggggtgccc gggctcagct cgggaggagg tgacctcagt cctctgacac ctgctccttc     1020
cccaaactct caggcagcgg ctccggccctg acggtgtggg gtacagggaa tccgcggagg     1080
cagttcatat actcgtctga cctggcccag ctctttatct gggtcctgcg ggagtacaat     1140
gaagtggagc ccatcatcct ctccgtgggc gaggaagatg aggtctccat caaggaggca     1200
gccgaggcgg tgggtggaggc catggacttc catggggaag tcacctttga tacaaccaag     1260
tcggatgggc agtttaagaa gacagccagt aacagcaagc tgaggaccta cctgcccgac     1320
ttccggttca cacccttcaa gcaggcgggtg aaggagacct gtgcttggtt cactgacaac     1380
tacgagcagg cccggaagtg aagctggaag acaggatcag gtgccagcgg accatcggct     1440
ggcagagccc agcggccacc acccgtcaac cctgccagga gctgagggca ccaccagca     1500
acctgggcct gcattccatc cgctctgcag cccaagcat ctttcagtg ggcccccat     1560
tcacgttggc cctcagggaa accagggctc ggggcaggcc cggcgctttg ctccccacac     1620
cagccccctg cgcgtgtcca ctctgacct gcacccact ccctgggagc caataaagt     1680
cattttcaca ggc                                     1693

```

<210> 16

<211> 2040

<212> DNA

<213> Homo sapien

<400> 16

```

gactccaagt cccagctgtg cccgggacgc tgcgcgcccc cgggactcca tttcccagag      60
tgccccgccc cagcctcccg gccccgcccc ccacctggct ccgcgggcag ccggagcgca     120
cccggccgga agccgctgtc ggggagccgg cgggtggggct ggacgcaggt gcaactgaca     180
tggtgaacc ccagggatcc atgcggattc tagtgacagg gggctctggg ctggtaggca     240
aagccatcca gaaggtggta gcagatggag ctggacttcc tggagaggac tgggtgtttg     300
tctcctctaa agacgccgat ctacaggata cagcacagac ccgcgccctg tttgagaagg     360
tccaaccac acacgtcatc catcttgctg caatgggtgg gggcctgttc cggaatatca     420
aatacaattt ggacttctgg aggaaaaacg tgcacatgaa cgacaacgtc ctgcactcgg     480
ccttcgaggt gggcgcccg cagggtgggt cctgcctgtc cacctgtatc ttccctgaca     540
agacgaccta ccgatagat gagaccatga tccacaatgg gcctccccac aacagcaatt     600
ttgggtactc gtatgccaaag aggatgatcg acgtgcagaa cagggcctac ttccagcagt     660

```



```

acggctgcac cttcaccgct gtcaccccca ccaacgtctt cgggccccac gacaacttca 720
acatcgagga tggccacgtg ctgcctggcc tcatccacaa ggtgcacctg gccaaagagca 780
gcggctcggc cctgacggtg tggggtagag ggaatccgcg gaggcagttc atatactcgc 840
tggacctggc ccagctcttt atctgggtcc tgcgggagta caatgaagtg gagcccatca 900
tcctctccgg tgggtacctc agccctcagc ccccttcctc catggtagga caggatccaa 960
ggctgtcctg ggaggctggc tgaggggtgg gagcaggaat attcatgggc cgtcgggtgg 1020
caccaagggg aggcctgacc cggtaggtcc ccacatgggc gaggaagatg aggtctccat 1080
caaggaggca gccgaggcgg tggtagggc catggacttc catggggaag tcaccgtatc 1140
ctctggctcc agtgctgggg gttgggcggg ggccccttgg gcgggctcag tggccacccg 1200
gtgtagggcc agcgtgggtg ggacactgcg gagggaggat gctcacctg gcaggcggag 1260
gagctctgct gtggctcccc tgggaatggc agagggtcaa gggcaagctg cgggggctgg 1320
ggtaggagtg ctctctcgc tgtggccttg acaagcctc catttgatac aaccaagtgc 1380
gatgggcagt ttaagaagac agccagtaac agcaagctga ggacctacct gcccgacttc 1440
cggttcacac ccttcaagca gggtagccc tgacccacac gccctccact tggtagggc 1500
ctgccccagc cctccctca tgetcgctac cagggcagag gtgaggtgcc ccctcagagc 1560
ctgtactgcc ctggaggtgg gacaccaaag ctctgagggg tgacttcctt gtccctgggc 1620
ctcctcccc cacttcctcc catgcggccc cgaccctccc tgtccatggc cctcacctgc 1680
ctgcctctgc agcgggtgaag gagacctgtg cttgggtcac tgacaactac gagcaggccc 1740
ggaagtgaag ctggaagaca ggatcaggtg ccagcggacc atcggctggc agagcccagc 1800
ggccaccacc cgtcaaccct gccaggagct gagggcacca cccagcaacc tgggcctgca 1860
ttccatccgc tctgcagccc caagcatctt tccagtgggg ccccatcca cgttggctct 1920
cagggaaacc agggtcgggg gcaggcccgg cgctttgctc cccacaccag cccctgcgc 1980
gtgtccactc tgatcctgca tccactccc tgggagccaa taaagtgcac ttccacaggc 2040

```

```

<210> 17
<211> 2481
<212> DNA
<213> Homo sapien

```

```

<400> 17
cgcccagcgc agccccggcc gctgggcgca cccgtcccgt tgcgtcccg gacgttgcct 60
ctacccccggg aacgtcgaga ctggagcgcc cgaactgagc caccttcgcg gacccccaga 120
gcggcggcgc gactcaccgc ggaggcgccc ggacgcgtca agtgcgctact ttcaaccgta 180
aatgggtgct gagcgtttag gggtaggggg caactcccta actcgtgggg ccctggccgg 240

```

ccctgccgcg	gcgtccgaga	ccttctgate	cctcgctgcg	cgcggcgcgg	ggttggcgac	300
ttgggctcag	gggcggcggg	gcggggaggt	gccttgacct	agggacccag	aggccgggag	360
caagtctcct	ttcgcttg	actgcagtgt	ccccagggc	aaggcctgaa	ccaggctcgcg	420
agggtctgtc	tcccgtctgg	ggagctgaac	tgcccagagc	actgtcgcag	tcgcggggac	480
ttgtctggag	cggcggggtc	tgagcccaag	tcggaagccg	ccgccgccgc	cagcgtcccg	540
tgcgctgcc	gccccgggac	cagccgcgaa	acttcttctc	gggacggaga	cagcgctcgc	600
cggccgggccc	gaggtaggag	ctggctgggc	tcggtagagc	gcgggcgaat	cgctttcggc	660
agcgaaggaa	cgtgtcggga	gaagccaact	gcaggaccag	gccccagcgg	gcgcgcagag	720
acgcctgcct	tgggcggagg	tccccccctc	cggttttctg	ggctctggta	gggccctctg	780
tcccgtgagc	tcctcactgg	acctgaggt	acaggcttca	gcaggccaag	cgcccgagt	840
tgcttgagg	attggcggcg	gggtggcatg	agaagggtgg	ccctcttctc	catctcccag	900
tcaggcgatt	tctctgctgt	ctggggcagg	ggcgcacacc	ccaggaaactt	actacgtaga	960
gggaagccag	cctggtctct	cccccttgcg	tcagtgtctc	taagctgccc	ctcaccactc	1020
tgggtagggg	tctttcaggt	gggctagggg	attctccttg	gaagtcccct	tgtttgctcc	1080
cctccatcag	ggcacctctg	agttagccaa	tcctcctctt	ttctggcctc	ttttttgcgg	1140
ggaggagagt	agagagagaa	gcctcctact	cggcaggtct	ctgagaggag	gtttgggggg	1200
gcttctgagg	ttgctcctcc	tgggcagagt	caaggggtgt	tgagggttcc	aacttttact	1260
ggacaatctc	tcaagctctg	attccctgat	ttctgggagg	tgtcagtgtg	gggcttgtgt	1320
cctgccccac	ccacacagca	tttgggtgcg	gcacatggga	tgggtcccca	agcccttttc	1380
ccacagttag	tttcagaaaa	tgcataagcc	acctaaggct	gggagagcct	ttctgtgctg	1440
gctgagatct	ggtggcatgt	gcatgcatat	gtgtgcaagt	gtgtgtgtgt	gtgtgcaagt	1500
gtgtgtgcgc	atgtgcatgt	gtgtggaaag	tatggtctgg	gtggctgtaa	gactggcccc	1560
tcttgctccc	tctgcaggcc	cagcccaata	ggaagcacct	gatccacact	ctttctagcc	1620
ccagggatgc	tcaggaaagc	agcctttctt	tagggctgtc	atctgttctg	gggatctgtt	1680
gctgtttctt	gggcccacag	aatacactcc	acacctcggg	actccataat	attggacaag	1740
gctggaaggc	ccatttggct	cctgcagcct	cttatgcagg	tgaggacaga	ccaagggtcc	1800
aggcaatgct	ggcctttgga	acacagaacc	tggtcactg	actccactc	cagttgtctt	1860
ttccctctgc	ctctggagca	accctctctc	ctccagagca	accctgtcc	ctccttcta	1920
cctctctcta	gggctgcacc	cccagctcac	cttagacctg	gcccttcta	cctgttgccc	1980
ctactgtctt	gtcccatccc	agtggtcagg	aggagagta	ctggccaaag	gcccagctcc	2040

33

```

acctgtgact taggggagtg ccaggcatca cccagaggtc ccggccctcg ggggcctggc 2100
aggctctgct gtgggggaag cagagttaga actggggctg cttctccact ggccgtttgt 2160
ctctgtcccc tccactggcc tcttgaagca cagagacct ccgggtactt tcttctctct 2220
gggcttccac tcatgctatt ccctgctccc ttttctacc tggagacccc cattcctagt 2280
cctcctcttc agcggctctc ccagagctgt cctcagcatg gcagcctcca cagcccctgg 2340
gtgtcacccc cggtttcgtg tctgcctcgg actccagacc tgcccctccc aggctggccc 2400
agatggatcc tctacagtga ctaacccaaa ccagcaccag gcagggaccc aataaacctt 2460
tgtgaatgga aaaaaaaaaa a 2481

```

```

<210> 18
<211> 2127
<212> DNA
<213> Homo sapien

```

```

<400> 18
aaaggggttcg ccgatcttgg ggagccacac agcccgggtc gcaggcacct ccccgccagc 60
tctcccgttt ctcgcacagc ttcccgacgc gtctgctgag ccccatggcc cagccacgc 120
tctccgcgcg ccccagcaat ccccggtccc tgcgggtggc gctgctgctc ctgctcctgg 180
tggccgccag cgggcgcgca gcagggtggg cccggcgccc tggggctccc gggccggagc 240
cggctggggg gggcgccccg cgccgacagc ccgctcaat cagcgagtct attcttcctt 300
aggagcgtcc gtggtcactg aactgcgctg ccagtgcctg cagacactgc aggggaattca 360
cctcaagaac atccaaagtg tgaatgtaag gtcccccgga cccactgcg cccaaaccga 420
agtcatgtaa gtcccgcccc gcgcttccgc tgccaccgct ggggtccccg actctccgcg 480
tgccccaacc ctgtccccag cccgacctcc tgtctcacgg aattcccttc tctctgcaga 540
gccacactca agaatgggaa gaaagcttgt ctcaaccccg catcccccat gggtcagaaa 600
atcatcgaaa agatactgaa caagtgagtt gtaatttcca tatacacagg cgactggagc 660
cgttggtcag aaatactggc atctgcccc taaaaataaa atcaggggaa cccaagagtt 720
agctgaagga ctagaaattg tgattattat ttccacaatt aaagttgcca ttaagggtat 780
taatctgctc tgggtgccaga ggatatttca gtatttcagt gtcttcatt tccaggccta 840
accctgggga gggtaaattg gggtttctag aatactaagt aaatttgact aacaatgcca 900
aatgcccc tgctgtttt gggacgttta cctcagtcct taatgggctt cagacctgaa 960
aatgcaatg tctggtttgg gcactttggg cctgtaagcc ttcttggtga agagcagaaa 1020
ctttgcagta acaccttcag tgagttcaag gctaggatcc ctgtcaacat tatttaattct 1080
ctttgagctt tagcaaaaata gggtagataa tggtagttat tcatagggt tgtgtgcagc 1140

```

34

ttaagggaga tcatgaatgg acatttctaa cccaatgtga caatcagtga ccaatatgtt 1200
 ttggcttttt caaaccaaga atctagttga gtccacagcc acttctttct tgaaaaaaaa 1260
 aaaaaaaaagc attcaggagt gcaagttttc tgctatagaa ggcagtgggt acagacagac 1320
 cttccctgag tggagggtctg ggagagagtc tgcatgggga catccctcca agcaacttca 1380
 gagtgacaga agagcagcct tccttagcac ctgcaacagt aacccttttc tcatcacagg 1440
 gggagcacca actgacagga gagaagtaag aagcttatca gcgtatcatt gacacttcct 1500
 gcagggtgggt ccctgccctt accagagctg aaaatgaaaa agagaacagc agctttctag 1560
 ggacagctgg aaaggactta atgtgtttga ctatttctta cgagggttct acttatttat 1620
 gtatttatatt ttgaaagctt gtattttaat attttacatg ctgttattta aagatgtgag 1680
 tgtgtttcat caaacatagc tcagtctga ttatttaatt ggaatatgat gggttttaaa 1740
 tgtgtcatta aactaatatt tagtgggaga ccataatgtg tcagccacct tgataaatga 1800
 cagggtgggg aactggaggg tggggggatt gaaatgcaag caattagtgg atcactgtta 1860
 gggtaaggga atgtatgtac acatctatct tttatacttt ttttttaaaa aaagaatgtc 1920
 agttgttatt tattcaaatt atctcacatt atgtgttcaa catttttatg ctgaagtttc 1980
 ccttagacat tttatgtctt gctttaggg cataatgcct tgtttaatgt ccattctgca 2040
 gcgtttctct ttccttgga aaagagaatt tatcattact gttacatttg tacaatatgac 2100
 atgataataa aagttttatg aaaaaaa 2127

<210> 19
 <211> 4323
 <212> DNA
 <213> Homo sapien

<400> 19
 gggacgtccg tggctccttg cgcacgtcgc agcgccctggc gcccggaag aggtggttgt 60
 gaggcagacg aactcgcggc tctccggctt ccgaggcttc cgagttgtcg gaggaagggg 120
 gcggcgagca ataagaaccg gccgcaccg gtccctcagcg actcttctga cctccgcgcg 180
 acgtacccgc cgccgccgtt ggctggagca tttgacattg tgcagcaaag aaatggttat 240
 ggagaagccc agtccgctgc ttgtagggag ggagtttgtg aggcaatatt atactttgct 300
 gaataaagct ccggaatatt tacacagggt ttatggcagg aattcttcct atgttcatgg 360
 tggagtagat gctagtggaa agccccagga agctgtttat ggccaaaatg atatacacca 420
 caaagtatta tctctgaact tcagtgaatg tcatactaaa attcgatcatg tggatgctca 480
 tgcaaccttg agtgatggag tagttgtcca ggtcatgggt ttgctgtcta acagtggaca 540
 accagaaaga aagtttatgc aaacctttgt tctggctcct gaaggatctg ttccaaataa 600

35

attttatggt cacaatgata tgtttcggtta tgaagatgaa gtgttatggt gattctgagc	660
ctgaacttga tgaagaatca gaagatgaag tagaagagga acaagaagaa agacaacat	720
ctcctgaacc tgtgcaagaa aatgctaaca gtggttacta tgaagctcac cctgtgacta	780
atggcataga ggagcctttg gaagaatcct ctcatgaacc tgaacctgag ccagaatctg	840
aaacaaagac tgaagagctg aaaccacaag tggaggagaa gaacttagaa gaactagagg	900
agaaatctac tactcctcct ccggcagaac ctgtttctct gccacaagaa ccaccaaaagg	960
ctttctcctg ggcttcagtg accagtaaaa acctgcctcc tagtggtact gtttcttct	1020
ctggaattcc acccatgtt aaagcaccag tctcacagcc aagagtcgaa gctaaaccag	1080
aagttcaatc tcagccacct cgtgtgcgtg aacaacgacc tagagaacga cctgggtttc	1140
ctcctagagg accaagacca ggcagaggag atatggaaca gaatgactct gacaaccgta	1200
gaataattcg ctatccagat agtcatcaac tttttgttgg taacttgcca catgatattg	1260
atgaaaatga gctaaaggaa ttcttcatga gttttggaaa cgttgtggaa cttcgcatca	1320
ataccaaggg tgttggggga aagcttccaa attttggtt tgtgggtttt gatgactctg	1380
aaccagttca gagaatctta attgcaaaac cgattatgtt tcgaggggaa gtacgtttta	1440
atgtggaaga gaaaaaaca agagctgcaa gagagcgaga aaccagaggt ggtggtgatg	1500
atcgcaggga tattaggcgc aatgatcgag gtcccggtgg tccacgtgga attgtgggtg	1560
gtggaatgat gcgtgatcgt gatggaagag gacctcctcc aaggggtggc atggcacaga	1620
aacttggtc tggaagagga accgggcaaa tggagggccg cttcacagga cagcgtcgct	1680
gaagctccac tgttgcaaa gtcttgagcag tggtagatta ttcatcgtgt ttgcattctt	1740
gttaattttt tttttggctt tggaatgtga cacagcctt ttgatcattt ctttgatgtg	1800
aaaagcatct ttggttatca gttaaattga ggtggacatt atttcccaa tttcacaaca	1860
ggattcacat tgtaattta taaatctaga cttggagaat taaggactga gaaatgacca	1920
tatcttaaac tatctacgac aaagtgaact taaaaggaca tgccactga attcaggtcc	1980
tttgagtaaa aaaaaaatct tctgctgcac attttggtta agtggtactg tttctgcctg	2040
ttaatgctgg gaacacaaat agtgcaattt gtgcaattgg agaactctgc ctttttctt	2100
ggctcccccc aaaaatacaa accaacagaa acttggtatg cactcatcaa aatgtactaa	2160
tgggtactct gaactcatta acattgacat ctgcaacagg aggcaacagg gaaaaaatct	2220
catcttcttt tccagtagaa aatagtttgt gaaatgatga gggcatttta tctgcttgct	2280
gtgaccagcg tgtgtacaca taaaccttaa caagactaca agtatattcc agaaggaaat	2340
cattttagtt atgaactaaa taataaaaat tagaacttca aatgcgatgg tcttgactat	2400
tagaccagat ttagtagctc catatctaag attttctac ctgcccctct tcagtacagg	2460

gatggctggc	tgctcaacac	actcctcctc	cccttttttc	ctttctttaa	gctgtgtaca	2520
gtgaaaattg	tctttactgt	atTTTTgttc	tctggtaatg	taataagcat	gatggtgcct	2580
tctattaata	catcattcca	gtcttgctgg	taattttgta	cagtatagtg	tatgaattgc	2640
tgtgctgcaa	agccaaacag	ctgcaaaatg	ttgaaaaatc	atcgaaatgt	ataaaaattg	2700
cagtatcttt	aaaatcagta	aaatggacta	gcatattatt	tatcttggtc	ttcagttaac	2760
aactttgtgt	tctctgtggg	agggaggagg	tcctgtgtgt	ttgtggggag	aggggaaggag	2820
gaagtcagtt	atttgagtaa	gcctctagtt	gacttttctc	ttagcctgaa	tgtggacggt	2880
gaaacatatc	acttcagggc	ttggaaaagt	cagtcaactt	gacgtacatt	tttagtgaca	2940
ttttaaagc	agtcagattc	tataaatggc	aagtaagcct	gaagtgagga	tactgcaatt	3000
ttcggagaaa	agaacagcag	ctctttaagt	gtttgcattt	tctatttggg	gggcagggaa	3060
ctgtcattca	ttttgcacaa	ttcttgaact	gatgtcagca	cccgagtggc	tcctgaattt	3120
aagtctggga	cgacatcttt	tattttttaca	tgaatcttta	aacaattctg	tgagcaaagt	3180
ttgtagctgc	tggattattg	tctgtcttta	tagcaagttc	cagtaaacca	caagtatggc	3240
aaagcttatc	caattttatg	cttggagcag	tcagtacata	ccagtttctg	atgtttcagg	3300
caggagtggg	gtaaataagt	gtgaccactt	aaagctgctc	gtagcatgg	aagacttctc	3360
cattctatct	ttgtaaaaca	gacaagatat	gcacttgaca	tagtagcaaa	ttggttctga	3420
attatgcaac	tgtttgctat	ttagtaaaact	agcaaagtat	gcatgtattt	tgtttttcat	3480
gtactgggca	atatgagtaa	aatctgtccc	tttttcccc	tttgaatgag	gtcttccatg	3540
tttgagggaa	agtcttgcac	tattgcatat	atTTTgggga	cacagatttt	catagtttcc	3600
atTTTTgggg	ggcttaagga	tttttttttt	ttctgtttga	aacagtttta	tactttctga	3660
tatagtactt	gaaattctta	ccagaaaatt	actttggagt	tttgaagcct	ttattaatac	3720
tacttttaaa	gaagcagttg	ttttattgtc	aatgtttttt	ttcccccaag	catattttct	3780
tgtatttctg	tttccatata	tatatatata	tataatttcc	aattcaggat	attgccctgc	3840
catccatgaa	aactgttctg	gcacccaaaag	taatgacaaa	tgttaagtgt	aataatagaa	3900
aagtagagca	aagagccatt	cagcttcagt	ctttacatac	catgaataaa	acattaaaac	3960
atcatatgga	gaagtttaca	tggtgattgt	tcacctgcag	tactgtggag	ttttaacatt	4020
ttgtcctctt	ttcagtgaaa	cagagtaaaa	atattcatct	accattactg	ttatttgctg	4080
atTTTgtttt	atTTTTtgat	ggtaatatcc	tatccttatg	acactattgc	aaccaaattg	4140
gctttaccat	cttggcttta	gtaggtatag	aagacaatgg	attaccatct	ttattgctgt	4200
aatgtgttaa	gcattatatg	ctagtagaat	ctagtttaat	tgtttcaggt	ggaaagtatt	4260

37

ctttgagttt ccatattgaa tgtgttttga ctaaacaac aataaactac tgatgtctgc 4320
 agc 4323

<210> 20
 <211> 3590
 <212> DNA
 <213> Homo sapien

<400> 20
 gactgggagg mamagagagt cagcttctgc tggcttccca atcagaagat gaagtagaag 60
 aggaacaaga agaaagacaa ccatctcctg aacctgtgca agaaaatgct aacagtgggt 120
 actatgaagc tcaccctgtg actaatggca tagaggagcc tttggaagaa tcctctcatg 180
 aacctgaacc tgagccagaa tctgaaacaa agactgaaga gctgaaacca caagtggagg 240
 agaagaactt agaagaacta gaggagaaat ctactactcc tcctccggca gaacctgttt 300
 ctctgccaca agaaccacca aagccaagag tcgaagctaa accagaagtt caatctcagc 360
 cacctcgtgt gcgtgaacaa cgacctagag aacgacctgg ttttcctcct agaggaccaa 420
 gaccaggcag aggagatatg gaacagaatg actctgacaa ccgtagaata attcgctatc 480
 cagatagtca tcaacttttt gttggttaact tgccacatga tattgatgaa aatgagctaa 540
 aggaattcctt catgagtttt ggaaacgttg tggaacttcg catcaatacc aagggtgttg 600
 ggggaaagct tccaaatttt ggttttgtgg tttttgatga ctctgaacca gttcagagaa 660
 tcttaattgc aaaaccgatt atgtttcgag gggaagtacg tttaaattgtg gaagagaaaa 720
 aaacaagagc tgcaagagag cgagaaacca gaggtggtgg tgatgatcgc agggatatta 780
 ggcgcaatga tcgaggtccc ggtggtccac gtggaattgt ggggtggtgga atgatgcgtg 840
 atcgtgatgg aagaggacct cctccaaggg gtggcatggc acagaaactt ggctctggaa 900
 gaggaaccgg gcaaatggag ggccgcttca caggacagcg tcgctgaagc tccactgttg 960
 gcaaagtctt ggcagtggta cattattcat cgtgtttgca ttcttgtaa tttttttttt 1020
 ggctttggaa tgtgacacag cctttttgat catttctttg atgtgaaaag catctttggt 1080
 tatcagttaa attgaggtgg acattatttc cccaatttca caacaggatt cacattgtta 1140
 atttataaat ctagacttgg agaattaagg actgagaaat gaccatatct taaactatct 1200
 acgacaaagt gaacttaaaa ggacatgcc actgaattca ggtcctttga gtaaaaaaaa 1260
 aatcttctgc tgcacatttt gtttaagtgt tactgtttct gcctgttaat gctgggaaca 1320
 caaatagtgc aatttgtgca attggagaat cttgcctttt ttcttggtc ccccaaaaaa 1380
 taaaaccaa cagaaacttg ttatgcactc atcaaatgt actaatgggt actctgaact 1440
 cattaacatt gacatctgca acaggaggca acagggaata aatctcatct tcttttccag 1500

38

tagaaaaatag tttgtgaaat gatgagggca ttttatctgc ttgctgtgac cagcgtgtgt	1560
acacataaac cttacaaga ctacaagtat attccagaag gaaatcattt tagttatgaa	1620
ctaaataata aaaattagaa cttcaaagtc gatggtcttg actatttagac cagatttagt	1680
agctccatat ctaagattttt tctacctgcc cctcttcagt acagggatgg ctggctgctc	1740
aacacactcc tcctcccctt ttttcctttc tttaagctgt gtacagtga aattgtcttt	1800
actgtatttt tgttctctgg taatgtaata agcatgatgg tgccttctat taatacatca	1860
ttccagtctt gctggtaatt ttgtacagta tagtgatga attgctgtgc tgcaaagcca	1920
aacagctgca aaatgttgaa aaatcatcga aatgtataaa aattgcagta tctttaaaat	1980
cagtaaaatg gactagcata ttattttatct tgttcttcag ttaacaactt tgtgttctct	2040
gtgggagggga gggagtcctg tgtgtttgtg gggagagggga aggaggaagt cagttatttg	2100
agtaagcctc tagttgactt ttctcttagc ctgaatgtgg acgttgaaac atatcacttc	2160
agggcttgga aaagtcagtc aacttgacgt acattttttag tgacatttta aaagcagtca	2220
gattctataa atggcaagta agcctgaagt gaggatactg caattttcgg agaaaagaac	2280
agcagctctt taagtgtttg cattttctat ttggggggca gggaaactgtc attcattttg	2340
cacaattctt gaactgatgt cagcacccga gtggctcctg aatttaagtc tgggacgaca	2400
tcttttatatt ttacatgaat ctttaaacaa ttctgtgagc aaagtttgta gctgctggat	2460
tattgtctgt ctttatagca agttccagta aaccacaagt atggcaaagc ttatccaatt	2520
ttatgcttgg agcagtcagt acataccagt ttctgatgtt tcaggcagga gtggggtaaa	2580
taagtgtgac cacttaaagc tgctcgttag catggaagac ttctccattc tatctttgta	2640
aaacagacaa gatatgcact tgacatagta gcaaattggg tctgaattat gcaactgttt	2700
gctatttagt aaactagcaa atgatgcatg tattttgttt ttcatgtact gggcaatatg	2760
agtaaaatct gtcccttttt ccccttttga atgaggtctt ccatgtttga gggaaagtct	2820
tgcactattg catatatattt ggggacacag attttcatag tttccatttt tggggggctt	2880
aaggattttt tttttttctg tttgaaacag ttttatactt tctgatatag tacttgaaat	2940
tcttaccaga aaattacttt ggagttttga agcctttatt aatactactt ttaaagaagc	3000
agttgtttta ttgtcaatgt tttttttccc ccaagcatat tttcttgat ttctgtttcc	3060
atatatatat atatataaa atttccaatt caggatattg ccctgccatc catgaaaact	3120
gttctggcac caaaagtaat gacaaatgtt aagtgtaata atagaaaagt agagcaaaga	3180
gccattcagc ttcagtcctt acataccatg aataaaacat taaaacatca tatggagaag	3240
tttacatggg gattgttcac ctgcagtact gtggagtttt aacattttgt cctcttttca	3300
gtgaaacaga gtaaaaatat tcatctacca ttactgttat ttgctgattt tgttttatatt	3360

ttgatggta atattctatc cttatgacac tattgcaacc aaattggctt taccatcttg 3420
 gcttagtag gtatagaaga caatggatta ccatctttat tgctgtaatg tgttaagcat 3480
 tatatgctag tagaatctag tttaattggt tcaggtggaa agtattcttt gagtttccat 3540
 attgaatgtg ttggactaa acaacaata aactactgat gtctgcagca 3590

<210> 21
 <211> 3200
 <212> DNA
 <213> Homo sapien

<400> 21
 aggtgtcta gaggatccct cgggccacca agacaggcag aggagatatg gaacagaatg 60
 actctgacaa ccgtagaata attcgctatc cagatagtca tcaacttttt gttggtaact 120
 tgccacatga tattgatgaa aatgagctaa aggaattctt catgagtttt ggaaacgttg 180
 tggaacttcg catcaatacc aagggtgttg ggggaaagct tccaaatttt ggttttgtgg 240
 tttttgatga ctctgaacca gttcagagaa tcttaattgc aaaaccgatt atgtttcgag 300
 gggaagtacg tttaaattgt gaagagaaaa aaacaagagc tgcaagagag cgagaaacca 360
 gaggtggtgg tgatgatcgc agggatatta ggcgcaatga tcgaggtccc ggtggtccac 420
 gtggaattgt ggggtgtgga atgatgcgtg atcgtgatgg aagaggacct cctccaaggg 480
 gtggcatggc acagaaactt ggctctggaa gaggaaccgg gcaaattggag ggccgcttca 540
 caggacagcg tcgctgaagc tccactgttg gcaagtctt ggcatggta cattattcat 600
 cgtgtttgca ttcttgtaa tttttttttt ggctttggaa tgtgacacag cttttttgat 660
 catttctttg atgtgaaaag catctttggt tatcagttaa attgaggtgg acattatttc 720
 cccaatttca caacaggatt cacattgtta atttataaat ctagacttgg agaattaagg 780
 actgagaaat gaccatatct taaactatct acgacaaagt gaacttaaaa ggacatgccc 840
 actgaattca ggtcctttga gtaaaaaaaaa aatcttctgc tgcacatttt gtttaagtgt 900
 tactgtttct gcctgttaat gctgggaaca caaatagtgc aatttgtgca attggagaat 960
 cttgcctttt ttcttggtc cccccaaaaa tacaaccaa cagaaacttg ttatgcactc 1020
 atcaaatgt actaatgggt actctgaact cattaacatt gacatctgca acaggaggca 1080
 acagggaaaa aatctcatct tcttttccag tagaaaatag tttgtgaaat gatgagggca 1140
 ttttatctgc ttgctgtgac cagcgtgtgt acacataaac cttaacaaga ctacaagtat 1200
 attccagaag gaaatcattt tagttatgaa ctaaataata aaaattagaa cttcaaatgc 1260
 gatggtcttg actattagac cagatttagt agctccatat ctaagatttt tctacctgcc 1320
 cctcttcagt acagggatgg ctggctgctc aacacactcc tcctccctt ttttcctttc 1380

tttaagctgt gtacagtga aattgtcttt actgtatttt tgttctctgg taatgtaata	1440
agcatgatgg tgccttctat taatacatca ttccagtctt gctggtaatt ttgtacagta	1500
tagtgtatga attgctgtgc tgcaaagcca aacagctgca aaatgttgaa aaatcatcga	1560
aatgtataaa aattgcagta tctttaaaat cagtaaaatg gactagcata ttatttatct	1620
tgttcttcag ttaacaactt tgtgttctct gtgggagggg gggagtcctg tgtgtttgtg	1680
gggagagggg aggaggaagt cagttatttg agtaagcctc tagttgactt ttctcttagc	1740
ctgaatgtgg acgttgaaac atatcacttc agggcttgga aaagtcagtc aacttgacgt	1800
acatttttag tgacatttta aaagcagtc gattctataa atggcaagta agcctgaagt	1860
gaggatactg caattttcgg agaaaagaac agcagctctt taagtgtttg cattttctat	1920
ttggggggca gggaaactgt attcattttg cacaattctt gaactgatgt cagcaccgga	1980
gtggctcctg aatttaagtc tgggacgaca tcttttattt ttacatgaat ctttaaaca	2040
ttctgtgagc aaagtttgta gctgctggat tattgtctgt ctttatagca agttccagta	2100
aaccacaagt atggcaaagc ttatccaatt ttatgcttgg agcagtcagt acataccagt	2160
ttctgatgtt tcaggcagga gtggggtaaa taagtgtgac cacttaaagc tgctcgtag	2220
catggaagac ttctccattc tatctttgta aaacagacaa gatatgcact tgacatagta	2280
gcaaattggg tctgaattat gcaactgttt gctatttagt aaactagcaa atgatgcatg	2340
tattttgttt ttcattgtact gggcaatatg agtaaaatct gtcccttttt cccctttga	2400
atgaggtctt ccatgtttga gggaaagtct tgcactattg catatatattt ggggacacag	2460
attttcatag ttccattttt tggggggcctt aaggattttt tttttttctg tttgaaacag	2520
ttttatactt tctgatatag tacttgaaat tcttaccaga aaattacttt ggagttttga	2580
agcctttatt aatactactt ttaaagaagc agttgtttta ttgtcaatgt tttttttccc	2640
ccaagcatat tttcttgat ttctgtttcc atatataat atatataaa atttccaatt	2700
caggatattg ccctgccatc catgaaaact gttctggcac caaaagtaat gacaaatgtt	2760
aagtgtaata atagaaaagt agagcaaaga gccattcagc ttcagtcttt acataccatg	2820
aataaaacat taaaacatca tatggagaag ttacatggg gattgttcac ctgcagtact	2880
gtggagtttt aacattttgt cctcttttca gtgaaacaga gtaaaaatat tcatctacca	2940
ttactgttat ttgctgattt tgttttattt tttgatggta atattctatc cttatgacac	3000
tattgcaacc aaattggcct taccatcttg gcttttagtag gtatagaaga caatggatta	3060
ccatctttat tgctgtaatg tgtaagcat tatatgctag tagaatctag ttttaattgtt	3120
tcagggtggaa agtattcttt gagtttccat attgaatgtg tttggactaa acaacaata	3180

41

aactactgat gtctgcagca

3200

<210> 22

<211> 627

<212> DNA

<213> Homo sapien

<400> 22

```

ggtaggctccc gagtaggctc agagcgaccg tggccggggcg gaagcgggctt tctggcctaa      60
gctttaacag gcttcgcctg tgcttctctgt ttctcttttt accaaggacc cgccaacatg      120
ggccgcgcttc gcaccaaaac cgtgaagaag gcggcccggg tcatcataga aaagtactac      180
acgcgccttg gcaacgactt ccacacgaac aagcgcgtgt gcgaggagat cgccattatc      240
cccagcaaaa agctccgcaa caagatagca ggttacgtca cgcattctgat gaagcgaatt      300
cagagaggcc cagtaagagg tatctccatc aagctgcagg aggaggagag agaaaggaga      360
gacaattatg ttcttgaggt ctcagccttg gatcaggaga ttattgaagt agatcctgac      420
actaaggaaa tgctgaagct tttggasttc ggcagtctgt scaaccttca ggtcattcat      480
cccaactgta ggctgagtga cctgaagggt ggacagactg cagttgggat gaatttcaaa      540
acgcctcggg gacctgtttg aattttttct gtagtgctgt attattttca ataaatctgg      600
gacaacagca aaaaaaaaaa aaaaaaa      627

```

<210> 23

<211> 3187

<212> DNA

<213> Homo sapien

<400> 23

```

ttgctcgagc ggccggcgcc agtgtgctgg aaaggcgcg gcttcggttcg taaggaaggg      60
ggcctaggcc cgggcctgcg gtggtggggg ttgctgcgcg ccggggggtcg ctctgctgt      120
gtcttccgct ccagcttcgc ccacttcccc ttgccagcgg ggtgggcgcg gagaagacct      180
gccggagcca tggaggacga agtgggtccgc tttgccaaga agatggacaa gatgggtgcag      240
aagaagaagg tgagcgcggc gagcgggcg ggccgggagg gaggcggggc cgcggtccgg      300
cccgaggggg ccgtcccctc cctccttctc cctcccctcc ccctgcagcc tcccgtctgg      360
gagccggggc cgccgcgccc gcgtcccccg ggaggggtgtc cgggcccggg gcgtcgggtc      420
ccgcggaac ctgcgttagg acccgctcgc tcttgatatt ccagtccctg ctgcgtggtg      480
ccgtgtaata aacctaatcg atttagggat tagggcacgc tctccgttga ctttaaactt      540
tcgtgctagt aaaaccttgg catttagctt aggaattgma gcgtagaaag tcctcagaac      600
ttgaaagaat ggagtcttcc gtgatccgtg gccaacatct agaaagagca tcaaataagag      660
gcttgatgc ggtattagac atcgggctgg taaaatgtaa aacttctgtg tcttattttt      720

```

gaagagggct	gaacgtccgc	ctggttatct	caactgtgct	acatgtttgc	ttagcgatta	780
aaaacgcggc	tggagcattg	gatttgctaa	aggagcttaa	gaatattcct	atgaccctgg	840
aattactgca	gtccacaaga	atcggaatgt	cagttaatgc	tattcgcaag	cagagtacag	900
atgaggaagt	tacatctttg	gcaaagtctc	tcatcaaata	ctggaaaaaa	ttattagatg	960
ggccatcaac	tgagaaagac	cttgacgaaa	agaagaaaga	acctgcaatt	acatcgcaga	1020
acagccctga	ggcaagagaa	gaaagtactt	ccagcggcaa	tgtaagcaac	agaaaggatg	1080
agacaaatgc	tcgagatact	tatgtttcat	cctttcctcg	ggcaccaagc	acttctgatt	1140
ctgtgcggtt	gaagtgtagg	gagatgcttg	ctgcagctct	tcgaacaggg	gatgactaca	1200
ttgcaattgg	agctgatgag	gaagaattag	gatctcaaata	tgaagaagct	atatatcaag	1260
aaataaggaa	tacagacatg	aaatacaaaa	atagagtacg	aagtaggata	tcaaatactta	1320
aagatgcaaa	aaatccaaat	ttaaggaaaa	atgtcctctg	tgggaatatt	cctcctgact	1380
tatttgctag	aatgacagca	gaggaaatgg	ctagtgatga	gctgaaagag	atgcggaaaa	1440
acttgaccaa	agaagccatc	agagagcatc	agatggccaa	gactgggtggg	accagactg	1500
acttgttcac	atgtggcaaa	tgtaaaaaga	agaattgcac	ttacacacag	gtacaaaccc	1560
gtagtgctga	tgaaccaatg	acaacatttg	ttgtctgtaa	tgaatgtgga	aatcgatgga	1620
agttctgttg	agttggaaga	attggcaaaa	tatctggacc	attaagaaaa	cggattttgt	1680
aactagcttt	aaactaggcc	aagcaactag	ttttcctgca	aatcaaattt	ttaaagcaac	1740
ttgggttaga	ctttgttttt	gacctaacat	cccttcctta	aatgccttct	gtagtttcag	1800
atcagtaggg	agaccatata	ataattttat	ggtacctgtt	tcaaaacata	ttttttctgt	1860
ttttataagt	aagttgatat	taattaaact	cttggcaata	tttcttcttt	cttaaaggaa	1920
aatatacctt	aacttttttt	cttttacact	gtgaaacata	cacagtagaa	attctgttac	1980
tctctgttat	taatacataa	atgaaaatac	atttttttcc	atattggcat	gtagctacaa	2040
atattaaagg	aggagaaaag	gtaatatata	tttaggttta	ccaaatatgg	tgtgtattca	2100
aataataact	gaccagctta	tctaaaatgt	acataatttt	gaggtagctt	atgaatttga	2160
ttttaattat	tatgttcaca	agcttggaat	attagatatt	attttgcac	tgtaactaac	2220
cgtgatcatc	atctcttgta	atctcttgta	catgtatatt	acttgttctt	aatagatttt	2280
tggaaacaag	actttattga	gatcagtttg	gttttcctgt	taatttacct	gtttgacttt	2340
ataatgtgtt	ttagttttgc	agaagaacac	tggtgtagtt	tagaaggctt	ttcataaatc	2400
ccctcatagg	caaagatgaa	aacttcccac	tatttttttc	ccctcttagg	aagacatact	2460
ggaaagaaaa	tgtttagcat	cttagtgtag	tatagctatt	gtaaacagtt	catgactaga	2520

43

ttttgattcg gaaatctata ctgaccaagg attaatctta aggattgtat aattcattaa	2580
agctgtggtc tttccatgtg gagactgata gaaaataatt ttgtcccaag tcttatttgc	2640
tgactttttc tgtcatgagt gagattgttg aacaaactga atatatgggc tatagcaagt	2700
agctttacag tacagatctt acaattaagt tttgcttttg ttaaagtgtg taccattttt	2760
tctgtttgga gtaagacaaa aattgttttg acatagggtc cctagggtac acttgctcta	2820
gcatacttta aaggccactg ttgcaaagtc tacattttat gctgaatctg cattctgtca	2880
ggcaccgta gaaagacctc agtacatgct ttgcactctc ctttgctccc tttttccaat	2940
ttcttattgc atatcatttt gttgtaatac agaaagcagc atttttaaat gtccgtgtta	3000
agaattggcc cactgggtacc aactcacctc tattttgtca gttcatagtt gaagattttg	3060
ttttatttca aaaacaaagt acatttttga aataatgttt cagaataaaa taatctcact	3120
tttaagtgat ccattttaaa atttgtaatt caataaagtt ttttttgttg ttaaacataa	3180
aaaaaaaa	3187

<210> 24
 <211> 2990
 <212> DNA
 <213> Homo sapien

<400> 24	
cccggcattc cgtgctcctt gggtccggcg ttggagctct ttggggccca gctttgcgga	60
cccgggagct cgggacgcag gcggggcttg tgctccgcgg gggcagggcg taggggtgggc	120
ctcctacctc ccctgatctc gcggtttgtt ccgtttcatt ggagcttccc ggaccgtgtg	180
ctcgacggtg ccctaggtgc cgtggggcca cacgcgagtc tgataagcac cctcccccg	240
aatcatgcgg tgctgtgagg cctagcgaag atgaagatag aatgcaaggt agaaagtgct	300
ggataccttt agaaagctgc aggactggtg cgatgggagt tgagacgtaa gaacctgccc	360
gtccgtaggg ctctggatgc tgctgaggcc cgaggccct atggcagatt tgaaaattca	420
cccttgtaga gtcattoctg cctttgagcg gactcccttt taagagatct caagagagcg	480
ttcgggtggag gccctgggtc tgcacagctc acctccctgg gaactgctcg cccgagcgtc	540
ggagccggcg ctggccccct gcagccgga ggttgagcc gcaggagccc cggaggccca	600
ggacacaggg ctcttgctct tgcagaatcc acaggtcttt cttgaggaaa tctgtagaca	660
gaactttgtg ctgcgttttt atctagggaa ggaacagaag agtgtcgtct cctagaaatc	720
tagcactgga gaaacgagga aaattcttcc agcgatggtc tccactcag agctgaggaa	780
gcttttctac tcagcagatg ctgtgtgttt tgatgttgac agcacggtca tcagagaaga	840
aggaatcgat gagctagcca aaatctgtgg cgttgaggac gcggtgtcag aaatgacacg	900

gcgagccatg	ggcggggag	tgcctttcaa	agctgctctc	acagagcgct	tagccctcat	960
ccagccctcc	agggagcagg	tgagagact	catagcagag	caacccccac	acctgacccc	1020
cggcataagg	gagctggtaa	gtcgcctaca	ggagcgaaat	gttcagggtt	tcctaatatc	1080
tgggtggcttt	aggagtattg	tagagcatgt	tgcttcaaag	ctcaatatcc	cagcaaccaa	1140
tgtatttgcc	aataggctga	aattctactt	taacggtgaa	tatgcagggt	ttgatgagac	1200
gcagccaaca	gctgaatctg	gtggaaaagg	aaaagtgatt	aaacttttaa	aggaaaaatt	1260
tcatttttaag	aaaataatca	tgattggaga	tggtgccaca	gatatggaag	cctgtcctcc	1320
tgctgatgct	ttcattggat	ttggaggaaa	tgtgatcagg	caacaagtca	aggataacgc	1380
caaatgggtat	atcactgatt	ttgtagagct	gctgggagaa	ctggaagaat	aacatccatt	1440
gtcgtacagc	tccaaacaac	ttcagatgaa	tttttacaag	ttatacagat	tgatactgtt	1500
tgcttacagt	tgctattac	aacttgctat	agaaagttgg	tacaaatgat	ctgtacttta	1560
aactacagtt	aggaatccta	gaagattgct	tttttttttt	ttttaactgt	agttccagta	1620
ttatatgatg	actattgatt	tcctggagag	gttttttttt	tttttgagac	agaatcttgc	1680
tctgttgccc	aggctggagt	gcagtggcgc	ggtctcggtc	caactgcaagc	tctgcctccc	1740
aggttcacgc	cattctcctg	cctcagcctc	ccgagtagct	gggactacag	gcgcccgcga	1800
ccacatccgg	ctaatttttt	gtatttttag	tagagacggg	gtttgaccgt	gttagccagg	1860
atggtcttga	tctcctgacc	ttgtgatccg	cctgcctcag	cctcccaaag	tgctgggatt	1920
acaggcttgg	gccacgcgc	ccagccaatg	tcctagagag	ttttgtgatc	tgaattcttt	1980
atgtatatatt	gtagctatat	ttcatacaaa	gtgctttaag	tgtggagagt	caattaaaca	2040
cctttactct	tagaaatacg	gattcggcag	ccttcagtga	atattgggtt	ctctttggta	2100
tgtcaataaa	agtttatccg	tatgtcagaa	cggatttgtg	gaattttcag	ttctgcctct	2160
tgcatatttg	aatacccaat	atctcaggta	ccgtgttgaa	gcataaatag	ccccaggca	2220
agaaacaggc	ttaagactct	actttgctta	gtgctgggta	caatgtcctc	accccatgag	2280
cccactccca	cttcctgaag	ccctgggacc	tggaaagtcc	ttttttttac	ttcttgggat	2340
ttgtcatata	tgaaatgtta	caaaagtttt	ttgttttttg	ttttcctcca	atataacaat	2400
agagtcttaa	acaaaagtac	tgattaaaaa	tggatatattg	ggctcaggcc	tgtaatctca	2460
gcactgtggg	aggacaagtt	aagaggatca	cttgaggaca	ggaattcgag	accagcctgg	2520
acaacagagc	aagaccctgt	ctctgcaaaa	aataaaaaaca	ttagctggat	gtggcggcac	2580
gcacctgtag	tcccagctac	ttggtgggct	gagatgggag	gatcccttga	actcagtagt	2640
tcaaggctgc	agtgagctgt	gactctgcca	ttgtactcca	gcctgggtga	cagtgaagacc	2700
ccgtctctta	aaaaaaaaaa	aaaaaaaaaa	gatttttgca	aaatccctta	aagttaaaaag	2760

45

aattaaaaac aggttttact aattcacaag ggccagcatt ttcccctaaa tgaattaaat 2820
 ttttaataact cctaagaggc tattataaac catttctttt ttataaatgc cttggaaata 2880
 ctccgaaagc caggtttatg tttgttttga gcagagccaa gcaagggaaa tttcagaaca 2940
 tgctttgaat gtgtgagatg ctataccctc cctggaggaa aaatgggggg 2990

<210> 25
 <211> 3153
 <212> DNA
 <213> Homo sapien

<400> 25
 cccggcattc cgtgctcctt gggtccggcg ttggagctct ttggggccca gctttgcgga 60
 cccgggagct cgggacgcag gcggggcttg tgctccgcgg gggcagggcg taggggtgggc 120
 ctccctacctc ccctgatctc gcgggttggt ccgtttcatt ggagcttccc ggaccgtgtg 180
 ctcgacgggtg ccctaggtgc cgtggggcca cacgcgagtc tgataagcac cctcccccg 240
 aatcatgcgg tgctgtgagg cctagcgaag atgaagatag aatgcaagg agaaagtgtc 300
 ggataccttt agaaagctgc aggactgggt cgatgggagt tgagacgtaa gaacctgccc 360
 gtccgtaggg ctctggatgc tgctgaggcc cgaggcccct atggcagatt tgaaaattca 420
 cccttgtaga gtcattcctg cctttgagcg gactcccttt taagtttaca gaagcacttg 480
 cagaactcat cagaagccac cccgcttctc agcagatctc aagagagcgt tcgggtggagg 540
 ccctgggtct gcacagctca cctccctggg aactgctcgc ccgagcgtcg gagccggcgc 600
 tggccccctg cagccggaag gttgcagccg caggagcccc ggaggcccag gacacaggac 660
 agagtctcgt tctgtcgcgc aggetggagt gcagtgggtg gatctcggct cacttcaacc 720
 tccgcctaca ggggttcaagg gattctcctg cctcagcctc ctgagtatct gggctcctgc 780
 tcttgacagaa tccacaggtc tttcttgagg aaatctgtag acagaacttt gtgctgcgtt 840
 tttatctagg gaaggaacag aagagtgtcg tctcctagaa atctagcact ggagaaacga 900
 ggaaaattct tccagcgatg gtctccact cagagctgag gaagcttttc tactcagcag 960
 atgctgtgtg ttttgatggt gacagcacgg tcatcagaga agaaggaatc gatgagctag 1020
 ccaaaatctg tggcggtgag gacgcgggtg cagaaatgac acggcgagcc atgggcgggg 1080
 cagtgccttt caaagctgct ctccagagc gcttagccct catccagccc tccagggagc 1140
 aggtgcagag actcatagca gagcaacccc cacacctgac ccccggcata agggagctgg 1200
 taagtcgcct acaggagcga aatgttcagg ttttccta atctgggtggc tttaggagta 1260
 ttgtagagca tgttgcttca aagctcaata tcccagcaac caatgtatgt gccaataggc 1320
 tgaaattcta ctttaacggt gaatatgcag gttttgatga gacgcagcca acagctgaat 1380

ctggtggaaa aggaaaagtg attaaacttt taaaggaaaa atttcatttt aagaaaataa	1440
tcatgattgg agatggtgcc acagatatgg aagcctgtcc tcctgctgat gctttcattg	1500
gatttggagg aaatgtgatc aggcaacaag tcaaggataa cgccaaatgg tatatcactg	1560
attttgtaga gctgctggga gaactggaag aataacatcc attgtcgtac agctccaaac	1620
aacttcagat gaatttttac aagttataca gattgatact gtttgcttac agttgcctat	1680
tacaacttgc tatagaaagt tggtagaaat gatctgtact ttaactaca gtttaggaatc	1740
ctagaagatt gctttttttt tttttttaac tgtagttcca gtattatatg atgactattg	1800
atttcctgga gaggtttttt ttttttttga gacagaatct tgctctgttg ccagggtgg	1860
agtgcagtgg cgcggtctcg gctcactgca agctctgcct ccagggttca cgccattctc	1920
ctgcctcagc ctcccagta gctgggacta caggcgcccg ccaccacatc cggctaattt	1980
tttgatattt tagtagagac ggggtttgac cgtgttagcc aggatggtct tgatctcctg	2040
accttgatcg ccgctgcct cagcctccca aagtgtggg attacaggct tgggccaccg	2100
cgccagcca atgtcctaga gagttttgtg atctgaattc tttatgtata tttgtagcta	2160
tatttcatac aaagtgcctt aagtgtggag agtcaattaa acacctttac tcttagaaat	2220
acggattcgg cagccttcag tgaatattgg tttctctttg gtatgtcaat aaaagtttat	2280
ccgtatgtca gaacggattt gtggaatttt cagttctgcc tcttgcatat tggaaatacc	2340
aatatttcag gtaccgtgtt gaagcataaa tagccccag gcaagaaaca ggcttaagac	2400
tctactttgc ttagtgctgg gtacaatgtc ctcaacctat gagccactc ccacttctg	2460
aagccctggg acctggaaag tccttttttt tacttcttgg gatttgtcat atatgaaatg	2520
ttacaaaagt tttttgtttt ttgttttct ccaatataac aatagagtct taaacaaaag	2580
tactgattaa aaatggatat ttgggctcag gcctgtaatc tcagcactgt gggaggacaa	2640
gttaagagga tcacttgagg acaggaattc gagaccagcc tggacaacag agcaagaccc	2700
tgtctctgca aaaaaataaa acattagctg gatgtggcgg cagcacctg tagtcccagc	2760
tacttggtgg gctgagatgg gaggatccct tgaactcagt agttcaaggc tgcagtgagc	2820
tgtgactctg ccattgtact ccagcctggg tgacagtgag acccgtctc ttaaaaaaaa	2880
aaaaaaaaaa aaagattttt gcaaaatccc ttaaagttaa aagaattaaa aacaggtttt	2940
actaattcac aagggccagc attttcccct aaatgaatta aattttaata actcctaaga	3000
ggctattata aaccatttct tttttataaa tgccttgga atactccgaa agccaggttt	3060
atgtttgttt tgagcagagc caagcaagg aaatttcaga acatgctttg aatgtgtgag	3120
atgctatacc ctccctggag gaaaaatggg ggg	3153

<210> 26
 <211> 2968
 <212> DNA
 <213> Homo sapien

<400> 26
 cccggcattc cgtgctcctt ggttcggcg ttggagctct ttggggccca gctttgcgga 60
 cccgggagct cgggacgcag gcggggcttg tgctccgcgg gggcagggcg taggggtgggc 120
 ctctacctc ccctgatctc gcggtttgtt ccgtttcatt ggagcttccc ggaccgtgtg 180
 ctgcacgggtg ccctaggtgc cgtggggcca caccgcagtc tgataagcac cctcccccg 240
 aatcatgcgg tgctgtgagg cctagcgaag atgaagatag aatgcaaggt agaaagtgtc 300
 ggataccttt agaaagctgc aggactgggt cgatgggagt tgagacgtaa gaacctgccc 360
 gtccgtaggg ctctggatgc tgctgaggcc cgaggccctt atggcagatt tgaaaattca 420
 cccttgtaga gtcattcctg cctttgagcg gactcccttt taagtttaca gaagcacttg 480
 cagaactcat cagaagccac ccgcttatac agcagatctc aagagagcgt tcggtggagg 540
 ccctgggtct gcacagctca cctccctggg aactgctcgc ccgagcgtcg gagccggcgc 600
 tggccccctg cagccggaag gttgcagccg caggagcccc ggaggcccag gacacaggat 660
 ctagggaagg aacagaagag tgctcgtctcc tagaaatcta gcactggaga aacgaggaaa 720
 attcttccag cgatggtctc cactcagag ctgaggaagc ttttctactc agcagatgct 780
 gtgtgttttg atgttgacag cacggctatc agagaagaag gaatcgatga gctagccaaa 840
 atctgtggcg ttgaggacgc ggtgtcagaa atgacacggc gagccatggg cggggcagtg 900
 cctttcaaag ctgctctcac agagcgctta gccctcatcc agccctccag ggagcagggtg 960
 cagagactca tagcagagca acccccacac ctgacccccg gcataagggga gctggtaagt 1020
 cgctacagg agcgaaatgt tcagggttttc ctaatatctg gtggcttttag gagtattgta 1080
 gagcatgttg cttcaaagct caatatccca gcaaccaatg tatttgccaa taggctgaaa 1140
 ttctacttta acggtgaata tgcagggtttt gatgagacgc agccaacagc tgaatctggt 1200
 ggaaaaggaa aagtgattaa acttttaaaag gaaaaatttc attttaagaa aataatcatg 1260
 attggagatg gtgccacaga tatggaagcc tgcctcctg ctgatgcttt cattggattt 1320
 ggaggaaatg tgatcaggca acaagtcaag gataacgcc aatggatat cactgatttt 1380
 gtagagctgc tgggagaact ggaagaataa catccattgt cgtacagctc caaacaactt 1440
 cagatgaatt ttacaagtt atacagattg atactgtttg cttacagttg cctattacaa 1500
 cttgctatag aaagttggta caaatgatct gtactttaaa ctacagttag gaatcctaga 1560
 agattgcttt tttttttttt ttaactgtag ttccagtatt atatgatgac tattgatttc 1620

48

```

ctggagaggt tttttttttt tttgagacag aatcttgctc tgttgcccag gctggagtgc 1680
agtggcgcggt tctcggtcca ctgcaagctc tgcctcccag gttcacgcca ttctcctgcc 1740
tcagcctccc gagtagctgg gactacaggc gcccgccacc acatccgggt aattttttgt 1800
attttttagta gagacgggggt ttgaccgtgt tagccaggat ggtcttgatc tcttgacctt 1860
gtgatccgcc tgcctcagcc tcccaaagtg ctgggattac aggcttgggc caccgcgccc 1920
agccaatgtc ctagagagtt ttgtgatctg aattctttat gtatatattgt agctatatatt 1980
catacaaagt gctttaagtg tggagagtca attaaacacc tttactctta gaaatacggg 2040
ttcggcagcc ttcagtgaat attggtttct ctttggtatg tcaataaaag tttatccgta 2100
tgtcagaacg gatttgtgga attttcagtt ctgcctcttg catattggaa tacccaatat 2160
ttcaggtacc gtgtgaagc ataaatagcc cccaggcaag aaacaggctt aagactctac 2220
tttgcttagt gctgggtaca atgtcctcac cccatgagcc cactcccact tctgaagcc 2280
ctgggacctg gaaagtcctt ttttttactt cttgggattt gtcatatatg aaatgttaca 2340
aaagtttttt gttttttgtt ttctccaat ataacaatag agtcttaaac aaaagtactg 2400
attaaaaatg gatatttggg ctcaggcctg taatctcagc actgtgggag gacaagttaa 2460
gaggatcact tgaggacagg aattcgagac cagcctggac aacagagcaa gaccctgtct 2520
ctgcaaaaaa taaaaacatt agctggatgt ggcggcacgc acctgtagtc ccagctactt 2580
gggtgggctga gatgggagga tcccttgaac tcagtagttc aaggctgcag tgagctgtga 2640
ctctgccatt gtactccagc ctgggtgaca gtgagacccc gtctcttaa aaaaaaaaaa 2700
aaaaaaaaaga tttttgcaa atcccttaa gttaaaagaa ttaaaaacag gttttactaa 2760
ttcacaaggg ccagcatttt cccctaaatg aattaaattt taataactcc taagaggcta 2820
ttataaacca tttctttttt ataaatgcct tggaaatact ccgaaagcca ggtttatggt 2880
tgttttgagc agagccaagc aagggaaatt tcagaacatg ctttgaatgt gtgagatgct 2940
ataccctccc tggaggaaaa atggggggg 2968

```

<210> 27

<211> 2664

<212> DNA

<213> Homo sapien

<400> 27

```

tatatcaact ttctgtgtca agatcttttag gtgaagaaca gttttctgtg ggatctgtca 60
acttcatgaa acagaccaat atccagaaaa atactaatac cagagataca agtaaaaaaa 120
caaaagacca gctcatcatc gacgctggtc agaaacattt tggggctact gtgtgcaagt 180
cttgtggtat gatataact gcttccaacc ctgaagatga aatgcagcat gtacagcatc 240

```

49

accacaggtt	tctggaagga	atcaaatatg	tgggttgga	gaaagaacgt	gtagtagcag	300
agttttggga	tgggaaaatc	gtgttggttc	tgccacatga	tccaagcttt	gctatcaaaa	360
aggtagaaga	tgtccaagaa	cttggtgata	atgaattggg	cttccagcaa	gttggtccta	420
aatgtccaaa	caaaataaaa	acttttcttt	ttatatctga	tgaaaagaga	gtagttgggt	480
gtttaattgc	agaacccatc	aaacaggcat	ttcgtgtcct	gtctgaacca	attggtccag	540
aatccccaa	ctctacggaa	tgtcctaggg	cttggaatg	ttcagatgta	ccagaacctg	600
cagtctgtgg	gataagtaga	atctgggttt	tcagactgaa	gagaagaaag	cgcattgcaa	660
gacgactgg	tgataccctc	aggaattgct	tcatgtttgg	ctgttttctc	agcactgatg	720
aaatagcatt	ttctgaccca	acaccagatg	gcaagttatt	tgcaaccaag	tactgcaaca	780
cccctaattt	cctcgtatat	aattttaata	gttaaagctg	atttcagtta	taaaggagtt	840
actatctgga	taagttcaaa	gagctcctta	ttataaaata	caaactattt	aatatcaaaa	900
taaaaaatac	cgagactcac	actcatacac	acacacacac	acacacacac	acacacatat	960
cacagttgtg	ttccttatga	gttgaaaagt	caggaataaa	tttgttgaaa	attatctggg	1020
gattcaaagg	aaaaatcttt	ggggtgattc	cctgattagc	actctgaatg	tttaattatg	1080
aaactttgta	gctataactg	gaaaattacc	tgactctttg	taagagtatt	aaatacaaag	1140
tgatttttct	ctagaaatgt	gacctggctc	tttataaagc	ccactcttag	accaggatta	1200
tctaattgcca	catcagaagc	aaacaggcaa	atttaaactt	gggcaagtaa	tttctgtgcc	1260
caatttgtaa	agggaattcc	tgaatttttt	ttttttttaa	tagaggcatg	ggtctcactg	1320
tgttgcccag	gctggtctga	aacttttggg	ctcaagcgat	cctcccaaaa	cgctgggatt	1380
acagtcatga	gccaccgtgc	ccagcctaatt	tctgacttcc	tctatacaga	gtcttcactt	1440
gataggcact	cgtctgtagt	aactcagttt	gaatatcttt	agaaaatggt	tagaatttat	1500
ttgtaacaag	atggtaagga	ataagattat	cccatatgca	tttctgtaga	gcagaatttg	1560
atagcttagt	gttcaatctt	tttgaaaata	aatgtttacc	tgtcatcaga	tttaattaaa	1620
attatactta	gtaattgcac	tattacttag	ttaatttttg	ttgtatggaa	atattggtag	1680
tactactttg	ggaacctgtt	actgacaatt	gatgtcatta	acaaaatgcc	tagttggatt	1740
agatgttttc	atcttctaata	tttttgcttg	tttaaaatgc	accttacttg	ttctgagata	1800
cctggcaaaa	gtctttacaa	aatgtatgg	aatagaacca	aggttagtaa	atatacatag	1860
gctgggtggat	gagagaccat	ggaactgtgt	aaatacactt	aaatgttcac	acatttttct	1920
agtgtaattc	ttggatactt	taaaaagcaa	aacattgttc	aaattgtttt	gattctgaaa	1980
aatcattcaa	ctgctaactg	gcaataagac	tctaggcaag	tcgttttcca	gattgtaatt	2040
atatgtagaa	actattcatc	tgcattcatt	ttatttgcct	gtaagttaac	atgtttccaa	2100

50

```

aattttaaag cctgggtccc caaaagaatg tggaagtatt aaaatgtatg taattatgca 2160
aacattttta tgctattttc tgcacttatt tcttttaaata attttattta aaatttttaa 2220
ttaacatttt gtttgcttaa tgcttttggt atgaatcaat taaaattcctt tattttatac 2280
aactaaatct gatttttaatt attttattat gaaaatagta tatgttaaat aaaattcaga 2340
taatacagaa atagagattg ccaaaagaag aggcttcccg gccggggcgcg gtgggtcacg 2400
cctgtaatcc cagcactttg ggaggccgag gcgggtggat catgagggtca ggagatcgag 2460
accatcctgg ctaacaagggt gaaaccccgt ctctactaaa aatacaaaaa attagccggg 2520
cgcggtggcg ggcgcctgta gtcccagcta ctcgaggaggc tgaggcagga gaatggcgtg 2580
aaccggggag gcggagcttg cagtgcgagc agattgcgcc actgcagtcc gcagtcgggc 2640
ctgggcgaca gagcgagact ccgt 2664

```

<210> 28

<211> 1757

<212> DNA

<213> Homo sapien

<400> 28

```

ctagcatcat aggtcaaaga cgttattaag tatcgactct gaaaaagcaa ctataaatac 60
tgcaaatacga ctaagccaga ggacgcaaata ataccctgcg tgttgtgtgc tggactgtgg 120
ttgagacttc atggtcgtgc tctggggtca tgggggcagc tgactgctca taattaagat 180
tgctgttgcc gctgcctgtg ttccaaagag aacacaagag gacagagcta aggtgtagga 240
atgcggaaga acttgcagaa actagagcta gtcttcggat tgaaatgaat taccaataaa 300
gggttgggaa atggatcaag gagaggggag cgttttgatg gtccctgctc cgagggttcag 360
cccatttgca gttgctgata accttgctcag gtcaccagat aaaggcaagt cttgtcaccc 420
tcaggcacct ggctgcgaag taataaaacg gttgccttat catctaaagg tggagtgaat 480
caaacctcag ccaggggtaa ttatagtaaa ctgaaattgc ctgtccttta caaaaaaaaa 540
ggcggaggga gagagaaagg aaaagggtt aaggcaggaa attgtgtttg tctttctcct 600
tggttttaaa aaagaaacga gcaaaacagc aaaaagtttt ccaggaggatg agaaaagcag 660
ccgggctaga cccccctcac cctgggttaa caggattgcc gacagatcag gccacgtga 720
tggtcaagaa aagaggctca agctgatgaa aatccaagaa atcttcttag gagccttcct 780
ttctaaatgt aatgactga gtcattgctt tgctaaggag ggaaggaagt tacaatcaca 840
ccaagtgact cattcattgc tgtagctcc ggcagttgga tgaaaagttt tagcaaaagc 900
agcaacaaat gtactttaaa tactagcaca gtaagggaat tcctgtcttt caggatgaat 960
gccatccaca caaaagagct cctgttgaca tctcatttac aatctcccc aggacacaga 1020

```

51

```

caagaccctt tcaataaatc cagctcagaa acacctattg ttcaaaatct tcaactcgcc 1080
acaggctacc accactccct atggctttgc aaaattaaag atttagaaga aggatggggc 1140
gggtggaagct atgaaaagag gcaggaaaaa agttcctttg acccgatgct gtcagagagc 1200
gtgcatgaag aagaaagtta atggatattc catttatata agaagcgcc taaagaaatgcc 1260
tgtgacgttc gtgaactagt gattgtgaat tccaaatttg atgccaaactt tatgtgtaaa 1320
gaagctaact cctgccaca tctgggctga atgaacagct gggactatgc ttaacccatt 1380
cccagcttat aaaagcccca tggcagctgc agtgaagcat cagaaaagta tagtaagaag 1440
aaactgaatt tgaagtggat tcttacaag gaaaaagaaa atcactattg taactatacc 1500
aaattactat gtatgtgagc tcccagcttt gacatcaagt cgttacagcc acgtaatggg 1560
attacgggtg atttttactt tattctttgg gcttctctga attttcctat attttgaaaa 1620
agcacctgca ttattttgta attataaaaa agtataaata agaattatac acactattat 1680
ttcttcaactg ctatgacatg actagattac aatttttttc tttttgagat gactttcggt 1740
cttgtttccc aggctgg 1757

```

<210> 29

<211> 1299

<212> DNA

<213> Homo sapien

<400> 29

```

caggacacct tgggccagct tggttttact ttagatttca ctgttggtccc accccacttc 60
ttccttcacc aacatgcaag ttctttcctt ccctgccagc cagatagaca gatgggagag 120
gcagtgcgag ccttcgttgt cagtagttct ttgatgtgaa aggggcagca cagtcattta 180
aacttgatcc aacctctttg catcttacag agttaaacag ctaaaagaag taaaataaga 240
aggcaatgct tgtggaatgt acagtgcata ttgacggcac acgcctcatt acgattcgcc 300
tgcttgcttc tcctgttcaa tegtctcttt ggaaggcagt ggatttttct cttgcgtctc 360
tgtcttcttc agtttcgact tatcgaattt ctgatctca gccatatcg atttgtcaga 420
catggttacg gaggagaagc ggagcgaggc acgcgagaac gagcgaagtc tggccttcgt 480
caagatggtt ggacaccacg tcgccttctt ggaagcagac gtgcttcagg ctgagcggga 540
ccatggggcc ttccttcagg ccctgcggag gtggctggga tccgcagggc tccctcctt 600
caggaacaag tcacctgcac cgggtgccgt gacgtacgag ctgcccacac tgtataggac 660
ggaggactat tttcctgtgg acgccgggga agcacagcac ccccccgca cctgcctcgt 720
gcctttgtga gctttgtggt cttcccatca ggaacgctgg aaagtgacat tgtgtacaca 780
ctgcagcttg ggggtttttt ctttgtattg ctgtttattt tatattttta aaatatttaa 840

```

```

aaaaatgtcg agatggggtc tcactatggt gtccagactg atctcaaact cctggggtca      900
agtgatccac ccaccttggc cttccaaagt ggtgggatta tgggcaggag cctccgtgcc      960
caggctgctg ccattttcaa atttcctccc tgcctcatgt gagaccacag ggtttggaga     1020
agcagttgga acccacgtgt ggtgatgcct cccacatcgg cctgcttggg gttctacagg     1080
ggttgaggga ccaggcctgg ccggggctga tggacagtgg ggactttcct tctctccatg     1140
atggctttgc aggggctcca tggtccttct ctctgtgatg ggtttttgca cgggggtgtgc     1200
tctgccactg tgggtgggtgg gtggatgctg cttctgttgc ctccagacct cgggtgccac     1260
agccttgagg atccttccaa taaagggtgc aagagctca                               1299

```

```

<210> 30
<211> 2456
<212> DNA
<213> Homo sapien

```

```

<400> 30
gggttcacaa gcgagagcag ggggcacagc ttccaccgt cctagccagg gtgagaccac      60
cctgcctgct ggtggagcca ggcgggccca gagccacct ggaagatgga ggctgcacgg     120
cactgcgtgg tccggggcag cccagggca gcagagcatt ccctggcctt ccctgctggt     180
gccagctcct taccacagag acgccgcgtg gaactcacta ctggcgatcg cggacgcccc     240
aggaaggcga gtggcacgag gtgtggcggg gctgccacag gcagtcacgg agcaccgcgt     300
ggtgccaggc taccctca gatacagcct cgtgggctgg ctcagtggcg gccctgactc     360
tccaaggctct gaggtctgct ctgtcttcca ttccatgact cacaggaggt tcaaggttac     420
ttccactgtg gcagctgcct ccctcctccc tctccaggat ggaaggaggt cctgctgtgc     480
aaacctgcct ggctgtcccc ttccgggaca cggacagggg gtttctctggc agttcctggt     540
cctcccttga gggctaaggg tcccccggt ctgtggccgc ctccagcaca tccaccaga     600
gtccctgggc gtgaacacag ccgctgggga cgatcacgc ctgcccagcg tgcagccctg     660
gggctgcgtc cttatctccg ctctcctgc ctctgccca gctgttcaac gtctacccat     720
ggctcggggc cctgctccag ctgcaccggc ccgtcctgcg caagatcgag gaggtccgtg     780
ccattctgag gacctcctg gaggcgcgga ggccccacgt gtgcccgggg gaccccggtg     840
gcagctatgt ggacgccctg atccagcagg gacaggggga tgaccccgag ggctgtttg     900
ctgaggccaa cgcggtggcc tgcaccctgg acatggtcat ggccgggacg gagacgacct     960
cggccacgct gcagtgggac gcacttctga tgggccggca cccggacgtg cagggtgaga    1020
ccccggcgcc tggcgagacg gctccttctg cccccggggg accccaggg acgagggatg    1080
gcgctgccac ccaagcggcc cacccttctg ccaggccgg gtgcaggagg agctagaccg    1140

```

cgtgctgggc cctgggcgga ctccccggct ggaggaccag caggctctgc cctacacaag 1200
 cgccgtgctc caccaggtgc agcggttcat caccgtcctg ccgcacgtgc cccgctgcac 1260
 cgcgcccgac acacagctgg gcggcttcct gctccccaag ggcacgcccg tgattcccct 1320
 gctgacctcg gtgctcctgg atgagacaca gtggcagacc ccaggccagt tcaaccccg 1380
 ccatttcctg gacgcgaatg ggcactttgt gaagcgggag gccttcctgc ctttctctgc 1440
 aggccgccgc gtctgtgttg gggagcgcct ggccaggacc gagctcttcc tgctgtttgc 1500
 cggcctcctg cagaggtacc gcctgctgcc cccgcctggc gtcagtccgg cctccctgga 1560
 caccacgccc gcccgggctt ttaccatgag gccgagggcc caggccctgt gtgcggtgcc 1620
 caggccctag gagtcccccc agccccagg tctcctgac cactcccctc ccagccctgg 1680
 gtccctccac cctctctcct cccacccac agctcggact gctctgggag ggccctgagg 1740
 actcccaccc tcaccccccac ccccacaggg tcagcaactg cttccgggta caccaggac 1800
 taccctgcc cgaccctgtg ggacccccac ccctctgatg ctgtctgcag ctcagtccct 1860
 gccagccccc aggagcgcct ccaggggcccc gccactctc ccacccctga agctgcactc 1920
 ccacccacct agtccccccc agggcccccc agcacctaca gctggggctg caggagagaca 1980
 acgggtggct gcatccagcc agagacaggc gcagggtgggt gtccctcagcg tgcgagccct 2040
 gcacccccca ggtcctggga ctctgcaga cccactcca tccccgtcc tggaacactt 2100
 cctgcagctg tgcttgagg cagtcggcct gcagtgccag actctgagcc aagccactgg 2160
 ggccatgctg atgactggtg caggagaggca aggccacat tctccttcag agacaggcac 2220
 tggcgccaga ggcttccttg gggcgggggg agggcacctc agcccctgaa gacaagcagc 2280
 actgcagtgg caaaaatgga aacactgacc cgggtgcggtg gctcatgcct gtaatcccag 2340
 cactttggga ggccgaggca ggcgaatcac gaggtcagga gttcgagacc agcctgcccc 2400
 acatggtgaa accctgtctc tactaaaaat acaaaaaaat taaccgagca tgggtgg 2456

<210> 31
 <211> 642
 <212> DNA
 <213> Homo sapien

<400> 31
 gggaaaggcc atgttcctc ctctgtgcca tgtcccacg tctcaggcct ttgtcaacga 60
 atccctcagc ccggacagag ggccgtggca tggctaattg cttggagaag cagctgagg 120
 gctgctttat ggcataaggt gagcgtcctt tgtcttgagg tcctctgggc aggggacact 180
 tgtgcctggt gccctggga ggaaggagag gagcatgtgc gggaaagagc aggaggccga 240
 ggtgggctga tcacgaggtc aggagtcca gaccagcctg gccaacatgg cacacatggc 300

54

acgcacatcg tggcgatgtc gcgtgccatg tgtctgatgc tgctgagcgt ggcgcgagca 360
 ttcttgttga tgggtggccg cacataggag gtggcaggct tccgctggcc ggatctccgc 420
 ttaatgacca ccacgacacc tttgccgtcg gctgccggct ccacgcccac agtcttgccg 480
 tgaatcagtc cgttgtagcg gaaggaattg cgggccctca agttattggg ctcaagtgtg 540
 tagtttttgg attatcattt taaaacatgg aaaactggag cagttccgca cgaccaatcc 600
 agttgctaata gcctgagcat cggtagcgag acatacgcca gc 642

<210> 32
 <211> 3907
 <212> DNA
 <213> Homo sapien

<400> 32
 gcgaattggg cctctagacg catgctctag caggcgggcg agtggtgctgg aaaggaccaa 60
 tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
 tttctgctca gccatgtcgc cggctccaga tgcagccccg gctcctgcgt cgatctccct 180
 gtttgacctc agcgcggatg ctccgggtctt tcagggcctg agcctggtga gccacgcgcc 240
 tggggagggt ctggcccggg ctccgcgtac ttcctgttca ggctcagggg agagagaaag 300
 cccagaaaga aagctactcc agggctcctat ggatatttca gagaagttat tttgttcaac 360
 ttgtgaccag accttccaga accaccaaga acagagggaa cattataagc ttgactggca 420
 tcggtttaac ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga 480
 aaagcagagc tccacaggag atctttccag catctcggga tcagaagact cagactcagc 540
 cagtgaggag gacttgcaga cactggatcg ggagagggct acatttgaga agttgagccg 600
 acccccaggc ttttaccctc atcgagttct tttccagaat gccagggcc agtttcttta 660
 tgcctaccgc tgtgtcctag gccctcatca ggatcccccga gaagaggcag aactgctgct 720
 acagaacctg caaagtagag gtcccagaga ctgcgtggtg ctcatggctg cagctgggca 780
 ctttgctggt gctatatttc aaggaagaga agtggtgaca caaaaactt ttcaccgcta 840
 tacggttcgg gccaaagcgg gcacagccca ggggcttcgg gatgcccag gtgggccatc 900
 aactctgctg ggagccaacc tgaggcgcta caatgaagcc aactatata aggatgttcg 960
 tgacctgctg gcagggccaa gctgggctaa ggcgctggag gaggtggta caatactgtt 1020
 gcgtgctccc cgctctggcc ggtctttgtt ctttggaggc aaggagcac ccctgcaaag 1080
 gggggatccc cgactttggg atatccccct cgccaccgc agaccacct tccaagagct 1140
 acagcgtgtg ctccataagc tgaccacttt gcatgtctat gaagaagacc ctcggaagc 1200
 agtcagactg cactcacctc agacacactg gaaaacagta agagaggaga gaaagaagcc 1260

tactgaggaa gaaataagaa agatctgcag ggatgaaaag gaagcgctgg ggcagaatga	1320
ggaatctccc aaacagggtt tgattactat ctggcaactg tcagatctga gtttctgtcc	1380
taaaaatgca ctggcaaatt ccctactctc ataactgacc catttcctta tttaaacaca	1440
caccacaggt gtctatctta gcaattactg agacatgttt agtgaccttt ccaaactctat	1500
tatctctttt gttatggatg tttccccaga gctaagaaaa tctttgtgtg gagcagatgc	1560
gtggatgtgt tgcattggatc tcacatagtt tatctttgat tggatgctgt ggtttgagggc	1620
tgagggtctg tcccatgctt tctctatttg tgccggcacc taaggctact catacggaga	1680
gtccagggtc ttgttcttgt ctgggtgag agcctcattt ctgcatgtga tataatggac	1740
tcagctcttg atgtgctact tttatctggg tgacaaaata aaaacaattt catttcagggc	1800
tttcattctt atgcatctgc ttctacctc tgagtactcc ctagatggat ttcacatgtc	1860
aggtttttcc ctaggttcag ggtcggaggg agaagatggc tttcaggtag agttggagct	1920
agtggagttg actgtgggga ctctggatct ttgtgagtct gaagtattgc ccaagcggag	1980
gaggagaaaa aggaataaga aggagaaaag ccgagaccag gaggctgggg cacatcggac	2040
tcttctccag caaactcaag aagaggagcc ttccacacag tcatcccagg cagttgctgc	2100
ccccttgggc cctttgctgg atgaggccaa agcccctggt cagccagagc tctggaatgc	2160
actgcttgct gcttgccgag ctggagatgt tggagtgtta aagctgcagc tagctcccag	2220
ccctgcagac cctagagttc tgtctctgct cagtgcctcc ttgggtctcc gtggctttac	2280
tctcctgcat gcagcagctg cagctggaag aggtcagtg gttcgtctgc tgctggaagc	2340
aggtgctgac ccactgtgc agtgagtaaa ggtcccatc ctgagtcatt ttgggtatag	2400
agaggataat ttggagggtta aagttggcat tggctacttg gcaatattcc cttggtctgg	2460
ttgggtgatg ttgggaaagg caggaggcaa gttaaaccac tttttggagt ttttgcacct	2520
agggactctc gggcccggcc acctatact gttgcggctg acaaataaac acgtaatgag	2580
ttccgaaggt tcatggagaa gaatccagat gcctacgatt acaacaaggc tcaggatcatc	2640
tggaaatagga aggacaagca gagtgggaag gcatcacaga tcctggctag atcctttcta	2700
cttcttgcat ttttaccaaa gataacttat agggctctat atttgaaatc ctccaggtgc	2760
caggaccatt gacaccagaa atggaggcac ggcaggctac acggaaaagg gagcagaagg	2820
cagcccggcg gcaacgggag gaacagcagc agaggcagca ggagcaggag gagcgtgaac	2880
gagaagagca gcggcgattt gccgccctca gtgaccgaga gaagagagct ctggctgcag	2940
agcgccgact cgctgccag ttgggagccc ctacctctc aatccctgac tctgcaatcg	3000
tcaatactcg acgtgctgg agttgtgggg catccctcca aggcctgact ccctttcact	3060

56

```

acctcgactt ctctttctgc tccacacgtt gcctccagga tcatcgccgt caggcagggga 3120
ggccctcttc ctgatctctt acagctctac ctggggccaa ctcagggacc tgagagggca 3180
cattcacagc agccctaggt tttttcttcc ccgtgaaacc agagatgatt tggaagatgg 3240
gggtgaagga cactcgggaa ctagggcaaa gacagggcta gaggtatgtg gagctgggtac 3300
tgtctctgga attttaatca caataaagtt tggcaaggaa tgtgtacttg tacttacatt 3360
cagaggcact gtggccttta gttccttagg gtcaccgtgg ccagcaccaa gggatcctgc 3420
ccactccatg tcccaggtcc aagggttact tttttttttt tttttttttt ttgagatgga 3480
gtgtctcact gtcagccagg ctggaagcga gactctgtct caaaaaacac aaacaaacaa 3540
acaaaagaaa tgtaaaatct cagatcctac ctcagaccta tagaatcaga atttgtattt 3600
taacaagatt cccaggtgat cctaggcacc ttcaaaattg agaagcatgg ctttacttga 3660
ctgggctccc tgaggacagg gagtgccttg ttcattctcaa atcctcagcc tctagcacag 3720
ggccgggtat atagtaggtt aaatctggat gaatcaatgt cataaccaa aggctaaaac 3780
aggttcagaa ccaggggaaa ccattggaaa atcttcaggg tgatccgggg tgtggaaggg 3840
aggaaccagg tcacaagaag gacctcttcc ctggagagtc ctgtttcagc aaaacagcca 3900
ttccctc 3907

```

```

<210> 33
<211> 3608
<212> DNA
<213> Homo sapien

```

```

<400> 33
gcgaattggg cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgctca gccatgtcgc cggctccaga tgcagccccg gtcctgcgt cgatctccct 180
gtttgacctc agcgcggatg ctccgggtctt tcagggcctg agcctggtga gccacgcgcc 240
tggggagggt ctggcccggg ctccgcgtac ttctgttca ggctcagggg agagagaaag 300
cccagaaaga aagctactcc agggctcctat ggatatttca gagaagttat tttgttcaac 360
ttgtgaccag accttcaga accaccaaga acagagggaa cattataagc ttgactggca 420
tcggtttaac ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga 480
aaagcagagc tccacaggag atctttccag catctcggga tcagaagact cagactcagc 540
cagtgaggag gacttgcaag cactggatcg ggagagggct acatttgaga agttgagccg 600
acccccaggc ttttaccctc atcgagttct tttccagaat gccaggggcc agtttcttta 660
tgcctaccgc tgtgtcctag gccctcatca ggatcccca gaagaggcag aactgctgct 720

```

57

acagaacctg caaagtagag gtcccagaga ctgcgtggtg ctcattggctg cagctgggca	780
ctttgctggt gctatatattc aaggaagaga agtgggtgaca cacaaaactt ttcaccgcta	840
tacgggttcgg gccaaagcggg gcacagccca ggggcttcgg gatgcccag gatgggccatc	900
acactctgct ggagccaacc tgaggcgcta caatgaagcc acactatata aggatgttcg	960
tgacctgctg gcagggccaa gctgggctaa ggcgctggag gaggctggta caatactgtt	1020
gcgtgctccc cgctctggcc ggtctttgtt ctttggaggc aaggagcac ccctgcaaag	1080
gggggatccc cgactttggg atatccccct cgccaccgc agaccacct tccaagagct	1140
acagcgtgtg ctccataagc tgaccacttt gcatgtctat gaagaagacc ctcggaagc	1200
agtcagactg cactcacctc agacacactg gaaaacagta agagaggaga gaaagaagcc	1260
tactgaggaa gaaataagaa agatctgcag ggatgaaaag gaagcgtgg ggcagaatga	1320
ggaatctccc aaacaggggt tgattactat ctggcaactg tcagatctga gtttctgtcc	1380
taaaaatgca ctggcaaatt ccctactctc ataactgacc catttctta tttaaacaca	1440
caccacagt gtctatctta gcaattactg agacatgttt agtgacctt ccaaatctat	1500
tatctctttt gttatggatg tttcccaga gctaagaaaa tctttgtgtg gagcagatgc	1560
gtggatgtgt tgcatggatc tcacatagtt tatctttgat tggatgctgt ggtttgaggc	1620
tgagggtgct tccatgctt tctctattg tgcgccacc taaggctact catcaggaga	1680
gtccagggac ttgttctgt ctgggctgag agcctcatt ctgcatgtga tataatggac	1740
tcagctcttg atgtgctact tttatctggg tgacaaaata aaaacaattt catttcaggc	1800
tttcattctt atgcatctgc ttctaccctc tgagtactcc ctagatggat ttcacatgtc	1860
aggtttttcc ctaggttcag ggtcggagg agaagatggc tttcaggtag agttggagct	1920
agtggagttg actgtgggga ctctggatct ttgtgagtct gaagtattgc ccaagcggag	1980
gaggagaaaa aggaataaga aggagaaaag ccgagaccag gaggctgggg cacatcggac	2040
tcttctccag caaactcaag aagaggagcc ttcacacag tcatcccagg cagttgctgc	2100
ccccttgggc cctttgctgg atgaggccaa agcccctgg cagccagagc tctggaatgc	2160
actgcttgct gcttgccgag ctggagatgt tggagtgtta aagctgcag tagctcccag	2220
ccctgcagac cctagagttc tgtctctgct cagtgcctcc ttgggctccg gtggctttac	2280
tctcctgcat gcagcagctg cagctggaag aggtcagtg gttcgtctgc tgctggaagc	2340
aggtgctgac ccactgtgc agtgagtaaa ggtcccatc ctgagtcatt ttgggtatag	2400
agaggataat ttggaggtta aagttggcat tggctacttg gcaatattcc cttggtctgg	2460
ttgggtgatg ttgggaaagg caggaggcaa gttaaaccct tttttggagt ttttgcacct	2520
agggactctc gggcccgcc accttatact gttgcggctg acaaatcaac acgtaatgag	2580

58

ttccgaaggt	tcatggagaa	gaatccagat	gcctacgatt	acaacaaggc	tcaggtgcca	2640
ggaccattga	caccagaaat	ggaggcacgg	caggctacac	ggaaaaggga	gcagaaggca	2700
gccccggcggc	aacgggagga	acagcagcag	aggcagcagg	agcaggagga	gcgtgaacga	2760
gaagagcagc	ggcgatttgc	cgcctcagt	gaccgagaga	agagagctct	ggctgcagag	2820
cgcgcactcg	ctgcccagtt	gggagcccct	acctctccaa	tccctgactc	tgcaatcgtc	2880
aatactcgac	gctgctggag	ttgtggggca	tccctccaag	gcctgactcc	ctttcactac	2940
ctcgacttct	ctttctgctc	cacacgttgc	ctccaggatc	atcgccgtca	ggcagggagg	3000
ccctcttcct	gatctcttac	agctctacct	ggggccaact	cagggacctg	agagggcaca	3060
ttcacagcag	ccctaggttt	tttcttcccc	gtgaaaccag	agatgatttg	gaagatgggg	3120
gtgaaggaca	ctcgggaaact	agggcaaaga	cagggctaga	ggtatgtgga	gctggtactg	3180
tctctggaat	tttaatcaca	ataaagtttg	gcaaggaatg	tgtacttgta	cttacattca	3240
gaggcactgt	ggccttttagt	tccttagggg	caccgtggcc	agcaccaagg	gatcctgccc	3300
actccatgtc	ccaggtccaa	gggttacttt	tttttttttt	tttttttttt	gagatggagt	3360
gtctcactgt	cagccaggct	ggagtgcagt	ggtacaacct	ccactcactg	caacctctgc	3420
ctcctggatt	caagcaattc	tcctgccctc	agcctcccaa	gtagctggga	ttagagggtgc	3480
ccaccatcac	gcccggctaa	tttttgattt	tttttgagat	ggagtgtctc	actgtcagcc	3540
aggctggagt	gcagtggcac	aatctcgggt	cactgcaacc	tctgcctccg	agattcaagc	3600
gattctcc						3608

<210> 34
 <211> 3550
 <212> DNA
 <213> Homo sapien

<400> 34	
gcgaattggg	cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag	gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgctca	gccatgtcgc cggctccaga tgcagccccg gctcctgcgt cgatctccct 180
gtttgacctc	agcgcggatg ctccggtctt tcagggcctg agcctggtga gccacgcgcc 240
tggggaggct	ctggccccggg ctcccggtac ttcctgttca ggctcagggg agagagaaag 300
cccagaaaga	aagctactcc agggctctat ggatatttca gagaagttat tttgttcaac 360
ttgtgaccag	accttccaga accaccaaga acagaggga cattataagc ttgactggca 420
tcggtttaac	ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga 480
aaagcagagc	tccacaggag atctttccag catctcggga tcagaagact cagactcagc 540

cagtgaggag gacttgcaga cactggatcg ggagagggct acatttgaga agttgagccg 600
acccccagggc ttttaccctc atcgagttct tttccagaat gcccagggcc agtttcttta 660
tgccctaccgc tgtgtcctag gccctcatca ggatccccca gaagaggcag aactgctgct 720
acagaacctg caaagtagag gtcccagaga ctgcgtggtg ctcatggctg cagctgggca 780
ctttgctggg gctatatattc aaggaagaga agtggtgaca cacaaaactt ttcaccgcta 840
tacggttcgg gccaaagcggg gcacagccca ggggcttcgg gatgcccag gtggggccatc 900
acactctgct ggagccaacc tgaggcgcta caatgaagcc acactatata aggatgttcg 960
tgacctgctg gcagggccaa gctgggctaa ggcgctggag gaggctggta caatactgtt 1020
gcgtgctccc cgctctggcc ggtctttgtt ctttgagggc aaggagcac ccctgcaaag 1080
gggggatccc cgactttggg atatccccct cgccaccgc agaccacct tccaagagct 1140
acagcgtgtg ctccataagc tgaccacttt gcatgtctat gaagaagacc ctcggaagc 1200
agtcagactg cactcacctc agacacactg gaaaacagta agagaggaga gaaagaagcc 1260
tactgaggaa gaaataagaa agatctgcag ggatgaaaag gaagcgtgg ggcagaatga 1320
ggaatctccc aaacagggtt tgattactat ctggcaactg tcagatctga gtttctgtcc 1380
taaaaatgca ctggcaaatt ccctactctc ataactgacc catttcctta tttaaacaca 1440
caccacagt gtctatctta gcaattactg agacatgttt agtgacctt ccaaactctat 1500
tatctctttt gttatggatg tttccccaga gctaagaaaa tctttgtgtg gagcagatgc 1560
gtggatgtgt tgcattggtc tcacatagtt tatctttgat tggatgctgt ggtttgaggc 1620
tgagggctgc ttccatgctt tctctattg tgcggcacc taaggctact catacggaga 1680
gtccagggac ttgttcttgt ctgggctgag agcctcattt ctgcatgtga tataatggac 1740
tcagctcttg atgtgctact tttatctggg tgacaaaata aaaacaattt catttcaggc 1800
tttcattctt atgcatctgc ttctaccctc tgagtactcc ctagatggat ttcacatgtc 1860
aggtttttcc ctaggttcag ggtcggaggg agaagatggc tttcaggtag agttggagct 1920
agtggagttg actgtgggga ctctggatct ttgtgagtct gaagtattgc ccaagcggag 1980
gaggagaaaa aggaataaga aggagaaaag ccgagaccag gaggctgggg cacatcggac 2040
tcttctccag caaactcaag aagaggagcc ttccacacag tcatcccagg cagttgctgc 2100
cccttggggc cttttgctgg atgaggccaa agcccctggc cagccagagc tctggaatgc 2160
actgcttgct gcttgccgag ctggagatgt tggagtgtc aagctgcagc tagctcccag 2220
ccctgcagac cctagagttc tgtctctgct cagtgcctcc ttgggctccg gtggctttac 2280
tctcctgcat gcagcagctg cagctggaag aggcctcagt gtctgtctgc tgctggaagc 2340

60

```

agggactctc gggccccggcc accttatact gttgcgggctg acaaatcaac acgtaatgag 2400
ttccgaaggt tcatggagaa gaatccagat gcctacgatt acaacaaggc tcagggtcatc 2460
tggaatagga aggacaagca gagtgggaag gcatcacaga tcctggctag atccttttcta 2520
cttcttgcag ttttaccaaa gataacttat aggggtcttat atttgaaatc ctccaggtgc 2580
caggaccatt gacaccagaa atggaggcac ggcaggctac acggaaaagg gagcagaagg 2640
cagccccggc gcaacgggag gaacagcagc agaggcagca ggagcaggag gagcgtgaac 2700
gagaagagca gcggcgattt gccgccctca gtgaccgaga gaagagagct ctggctgcag 2760
agcgccgact cgctgcccag ttgggagccc ctacctctcc aatccctgac tctgcaatcg 2820
tcaatactcg acgctgctgg agttgtgggg catccctcca aggcctgact cccttttact 2880
acctcgactt ctcttttctgc tccacacgtt gcctccagga tcatcgccgt caggcagggga 2940
ggccctcttc ctgatctctt acagctctac ctggggccaa ctgagggacc tgagaggggca 3000
cattcacagc agccctaggt tttttcttcc ccgtgaaacc agagatgatt tggaagatgg 3060
gggtgaagga cactcgggaa ctagggcaaa gacagggcta gaggtatgtg gagctggtag 3120
tgtctctgga attttaata caataaagtt tggcaaggaa tgtgtacttg tacttacatt 3180
cagaggcact gtggccttta gttccttagg gtcaccgtgg ccagcaccaa gggatcctgc 3240
ccactccatg tcccagggtc aagggttact tttttttttt tttttttttt ttgagatgga 3300
gtgtctcact gtcagccagg ctggagtgcg gtggtacaac ctccactcac tgcaacctct 3360
gcctcctgga ttcaagcaat tctcctgccc tcagcctccc aagtagctgg gattagaggt 3420
gcccaccatc acgcccggct aatttttgta tttttttgag atggagtgtc tcaactgtcag 3480
ccaggctgga gtgcagtggc acaatctcgg ctcaactgaa cctctgcctc cgagattcaa 3540
gcgattctcc 3550

```

```

<210> 35
<211> 3562
<212> DNA
<213> Homo sapien

```

```

<400> 35
gcgaattggg cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgctca gccatgtcgc cggctccaga tgcagccccg gctcctgcgt cgatctccct 180
gtttgacctc agcgcggtatg ctccgggtctt tcagggcctg agcctgaggg aacattataa 240
gcttgactgg catcggttta acctaaagca acgtctcaag gacaagcctc tcctgtctgc 300
cctggacttt gaaaagcaga gctccacagg agatctttcc agcatctcgg gatcagaaga 360

```

61

ctcagactca gccagtgagg aggacttgca gacactggat cgggagaggg ctacatttga	420
gaagttgagc cgacccccag gcttttacct tcatcgagtt cttttccaga atgcccaggg	480
ccagtttctt tatgcctacc gctgtgtcct agggccctcat caggatcccc cagaagaggg	540
agaactgctg ctacagaacc tgcaaagtag aggtcccaga gactgcgtag tgctcatggc	600
tgcagctggg cactttgctg gtgctatatt tcaaggaaga gaagtggtag cacacaaaac	660
ttttcacgc tatacgggtc gggccaagcg gggcacagcc caggggcttc gggatgccc	720
aggtgggcca tcacactctg ctggagccaa cctgaggcgc tacaatgaag ccacactata	780
taaggatgtt cgtgacctgc tggcagggcc aagctgggct aaggcgctgg aggaggctgg	840
tacaatactg ttgcgtgctc cccgctctgg ccggtctttg ttctttggag gcaaggagc	900
acctctgcaa aggggggagc cccgactttg ggatatcccc ctgcccaccc gcagaccac	960
cttccaagag ctacagcgtg tgctccataa gctgaccact ttgcatgtct atgaagaaga	1020
ccctcgggaa gcagtcagac tgactcacc tcagacacac tggaaaacag taagagagga	1080
gagaaagaag cctactgagg aagaaataag aaagatctgc agggatgaaa aggaagcgct	1140
ggggcagaat gaggaatctc ccaaacaggg ttgtattact atctggcaac tgtcagatct	1200
gagtttctgt cctaaaaatg cactggcaaa ttccctactc tcataactga cccatttctt	1260
tatttaaaca cacaccaca gtgtctatct tagcaattac tgagacatgt ttagtgacct	1320
ttccaaatct attatctctt ttgttatgga tgtttcccca gagctaagaa aatctttgtg	1380
tggagcagat gcgtggatgt gttgcatgga tctcacatag ttatctttg attggatgct	1440
gtggtttgag gctgagggct gcttccatgc ttccctctat tgtgcggcac cctaaggcta	1500
ctcatacggg gagtccaggg acttggtctt gtctgggctg agagcctcat ttctgcatgt	1560
gatataatgg actcagctct tgatgtgcta cttttatctg ggtgacaaaa taaaaacaat	1620
ttcatttcag gctttcattc ttatgcatct gcttctacct tctgagtact ccctagatgg	1680
atttcacatg tcagggtttt ccctagggtc agggtcggag ggagaagatg gctttcaggt	1740
agagttggag ctagtggagt tgactgtggg gactctggat ctttgtgagt ctgaagtatt	1800
gccaagcgg aggaggagaa aaaggaataa gaaggagaaa agccgagacc aggaggctgg	1860
ggcacatcgg actcttctcc agcaaactca agaagaggag ccttccacac agtcatccca	1920
ggcagttgct gcccccttg gccctttgct ggatgaggcc aaagcccctg gtcagccaga	1980
gctctggaat gcaactgttg ctgcttgccg agctggagat gttggagtgc taaagctgca	2040
gctagctccc agccctgcag accctagagt tctgtctctg ctcaagtccc ccttgggctc	2100
cgggtggctt actctcctgc atgcagcagc tgcagctgga agaggctcag tggttcgtct	2160
gctgctggaa gcagggtgctg accccactgt gcagttagta aagggtccca tctgagtca	2220

```

ttttgggtat agagaggata atttggaggt taaagttggc attggctact tggcaatatt 2280
cccttggtct gggtgggtga tgttgggaaa ggcaggaggc aagttaaacc catttttgga 2340
gtttttgcac ctagggactc tcggggcccg ccaccttata ctgttgccgc tgacaaatca 2400
acacgtaatg agttccgaag gttcatggag aagaatccag atgcctacga ttacaacaag 2460
gctcaggtca tctggaatag gaaggacaag cagagtggga aggcatacaca gatcctggct 2520
agatcccttc tacttcttgc agttttacca aagataactt ataggtctt atatttgaaa 2580
tcctccaggt gccaggacca ttgacaccag aaatggaggc acggcaggct acacggaaaa 2640
gggagcagaa ggcagcccgg cggcaacggg aggaacagca gcagaggcag caggagcagg 2700
aggagcgtga acgagaagag cagcggcgat ttgccgccct cagtgaccga gagaagagag 2760
ctctggctgc agagcgccga ctcgctgccc agttgggagc ccctacctct ccaatccctg 2820
actctgcaat cgtcaatact cgacgctgct ggagttgtgg ggcatccctc caaggcctga 2880
ctccctttca ctacctgac ttctctttct gctccacacg ttgcctccag gatcatcgcc 2940
gtcaggcagg gaggccctct tcctgatctc ttacagctct acctggggcc aactcaggga 3000
cctgagaggg cacattcaca gcagccctag gttttttctt ccccgtagaa ccagagatga 3060
tttggaagat ggggggtgaag gacactcggg aactagggca aagacagggc tagaggtatg 3120
tggagctggg actgtctctg gaattttaat cacaataaag tttggcaagg aatgtgtact 3180
tgtacttaca ttcagaggca ctgtggcctt tagttcctta gggtcaccgt ggccagcacc 3240
aagggatcct gccactcca tgtcccaggt ccaagggtta cttttttttt tttttttttt 3300
ttttgagatg gagtgtctca ctgtcagcca ggctggagtg cagtgggtaca acctccactc 3360
actgcaacct ctgcctcctg gattcaagca attctcctgc cctcagcctc ccaagtagct 3420
gggattagag gtgcccacca tcacgcccg ctaatttttg tatttttttg agatggagtg 3480
tctcactgtc agccaggctg gagtgcagtg gcacaatctc ggctcactgc aacctctgcc 3540
tccgagattc aagcgattct cc 3562

```

```

<210> 36
<211> 4502
<212> DNA
<213> Homo sapien

```

```

<400> 36
gcgaattggg cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgctca gccatgtcgc cggctccaga tgcagccccg gctcctgcgt cgatctccct 180
gtttgacctc agcgcggatg ctccggtctt tcagggcctg agcctggtga gccacgcgcc 240

```


tggggaggct ctggcccggg ctccgcgtac ttctgttca ggctcagggg agagagaaag	300
cccagaaaga aagctactcc agggtcctat ggatatttca gagaagttat ttgtttcaac	360
ttgtgaccag accttccaga accaccaaga acagagggaa cattataagc ttgactggca	420
tcggtttaac ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga	480
aaagcagagc tccacaggag atctttccag catctcggga tcagaagact cagactcagc	540
cagtgaggag gacttgca cactggatcg ggagagggct acatttgaga agttgagccg	600
acccccaggc ttttaccctc atcgagtctt tttccagaat gccaggggcc agtttcttta	660
tgcctaccgc tgtgtcctag gccctcatca gaggcaagtg acagtacagg ttgcatggct	720
aacccagcc ttttgtacac ccagcctaga ttcccttag cctagtgcac ttctccctgc	780
ttctcttctt tcctgttggg tgtgggtatc acgcttgaca acatagtctt gaattgaaac	840
tggcctgaat tgtggaaggg ttaagtgtaa tggaattaag tttggaaaca aaccaattta	900
ggtttgaatt gtgattctac cacttcatct tgactgcctc gccagcctc aaatttcctt	960
gtttgtcaaa tggaaacaag cctcaggggtg gtaatattaa atgagataat gtttgtgaaa	1020
attggccctt agtttcttgc tctcagtga tgcaacagat atgtttctga tgcttctctg	1080
tcatcaagga tccccagaa gaggcagaac tgctgctaca gaacctgcaa agtagaggctc	1140
ccagagactg cgtgggtgctc atggctgcag ctgggcactt tgctgggtgct atatttcaag	1200
ggtgagaggg tgctgcatgg gggtaaggat ggagtggatc ctgttagcca gggagcacca	1260
gctgggtctc agtactgagt ctgtgctgtc tacagaagag aagtgggtgac acacaaaact	1320
tttcaccgct atacggttcg ggccaagcgg ggcacagccc aggggcttcg ggatgcccga	1380
ggtgggcat cactctctgc tggagccaac ctgaggcgct acaatgaagc cactatat	1440
aaggtgagtt aagcctttta gatctgggtg tactgatcca tacttttttt gcatctctgt	1500
tcaactcact gttaagtagg gaggtttttg actcatatcc ttccaatttt cattataggt	1560
agacaaagg agatgaggga ttccccagga gccttagtcc tctattcttt ctccgtgtta	1620
tcctctactt tccaccctt gcacttctgt gggccatttt ttttagttgg agatttataa	1680
atagcaagtg aaggtatgca ggactgagtt actctcttct ctaggatgtt cgtgacctgc	1740
tggcagggcc aagctgggct aaggcgctgg aggaggctgg tacaatactg ttgcgtgctc	1800
ccgctctgg ccggtctttt ttcttttgag gcaagggagc acccctgcaa agggggggtc	1860
ccgactttg ggatatcccc ctgcgccccc gcagaccac cttccaagag ctacagcgtg	1920
tgtccataa gctgaccact ttgcatgtct atgaagaaga ccctcgggaa gcagtcagac	1980
tgcactcacc tcagacacac tggaaaacag taagagagga gagaaagaag cctactgagg	2040

64

aagaaataag aaagatctgc agggatgaaa aggaagcgct ggggcagaat gaggaatctc	2100
ccaaacaggg tttgattact atctggcaac tgtcagatct gagtttctgt cctaaaaatg	2160
cactggcaaa ttccctactc tcataactga cccatttcct tatttaaaca cacaccaca	2220
gtgtctatct tagcaattac tgagacatgt ttagtgacct ttccaaatct attatctctt	2280
ttgttatgga tgtttcccca gagctaagaa aatctttgtg tggagcagat gcgtggatgt	2340
gttgcatgga tctcacatag tttatctttg attggatgct gtggtttgag gctgagggtc	2400
gcttccatgc tttcctctat tgtgcggcac cctaaggcta ctcatacgga gagtcagggt	2460
acttgttctt gtctgggctg agagcctcat ttctgcatgt gatataatgg actcagctct	2520
tgatgtgcta cttttatctg ggtgacaaaa taaaaacaat ttcatttcag gctttcattc	2580
ttatgcatct gcttctaccc totgagtact ccctagatgg atttcacatg tcagggtttt	2640
ccctagggtc agggtcggag ggagaagatg gctttcaggt agagttggag ctagtggagt	2700
tgactgtggg gactctggat ctttgtgagt ctgaagtatt gcccaagcgg aggaggagaa	2760
aaaggaataa gaaggagaaa agccgagacc aggaggctgg ggcacatcgg actcttctcc	2820
agcaaactca agaaggagg ccttccacac agtcatccca ggcagttgct gcccccttgg	2880
gcccccttgc ggatgaggcc aaagcccctg gtcagccaga gctctggaat gactgcttg	2940
ctgcttgccg agctggagat gttggagtgc taaagctgca gctagctccc agccctgcag	3000
accctagagt tctgtctctg ctcagtgtccc ccttgggctc cgggtggcttt actctcctgc	3060
atgcagcagc tgcagctgga agaggctcag tggttcgtct gctgctggaa gcaggtgctg	3120
acccactgt gcagtgagta aagggtcccca tcctgagtca ttttgggtat agagaggata	3180
atgttgagggt. taaagttggc attggctact tggcaatatt cccttggctt ggttgggtga	3240
tgttgggaaa ggcaggaggc aagttaaacc catttttgga gtttttgac ctagggactc	3300
tcgggcccgg ccaccttata ctgttgccgc tgacaaatca acacgtaatg agttccgaag	3360
gttcatggag aagaatccag atgcctacga ttacaacaag gctcaggtca tctggaatag	3420
gaaggacaag cagagtggga aggcacaca gatcctggct agatcctttc tacttcttgc	3480
agttttacca aagataactt atagggtctt atatttgaaa tcctccaggt gccaggacca	3540
ttgacaccag aaatggaggc acggcaggct acacggaaaa gggagcagaa ggcagcccgg	3600
cggcaacggg aggaacagca gcagaggcag caggagcagg aggagcgtga acgagaagag	3660
cagcggcgat ttgccgcctc cagtgaccga gagaagagag ctctggctgc agagcgccga	3720
ctcgctgcc agttgggagc ccctacctct ccaatccctg actctgcaat cgtcaatact	3780
cgacgtgct ggagttgtgg ggcacccctc caaggcctga ctccctttca ctacctgac	3840
ttctctttct gctccacacg ttgcctccag gatcatcgcc gtcaggcagg gaggcctct	3900

65

```

tcctgatctc ttacagctct acctggggcc aactcagggc cctgagaggg cacattcaca 3960
gcagccctag gttttttctt ccccgtagaa ccagagatga tttggaagat ggggggtgaag 4020
gacactcggg aactagggca aagacagggc tagaggtatg tggagctggg actgtctctg 4080
gaattttaat cacaataaag tttggcaagg aatgtgtact tgtacttaca ttcagaggca 4140
ctgtggcctt tagttcctta gggtcaccgt ggccagcacc aagggatcct gcccaactcca 4200
tgtcccaggt ccaaggggta cttttttttt tttttttttt ttttgagatg gagtgtctca 4260
ctgtcagcca ggctggagtg cagtggtaga acctccactc actgcaacct ctgcctcctg 4320
gattcaagca attctcctgc cctcagcctc ccaagtagct gggattagag gtgcccacca 4380
tcacgccccg ctaatttttg tatttttttg agatggagtg tctcactgtc agccaggctg 4440
gagtgcagtg gcacaatctc ggctcactgc aacctctgcc tccgagattc aagcgattct 4500
cc 4502

```

```

<210> 37
<211> 3316
<212> DNA
<213> Homo sapien

```

```

<400> 37
gcgaattggg cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgctca gccatgtcgc cggctccaga tgcagccccg gtcctgcgt cgatctccct 180
gtttgacctc agcgcggtg ctccgggtctt tcagggcctg agcctgggtga gccacgcgcc 240
tggggaggct ctggcccggg ctccgcgtac ttctgttca ggctcagggg agagagaaag 300
cccagaaaga aagctactcc agggctcctat ggatatttca gagaagttat tttgttcaac 360
ttgtgaccag accttcaga accaccaaga acagagggaa cattataagc ttgactggca 420
tcggtttaac ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga 480
aaagcagagc tccacaggag atctttccag catctcggga tcagaagact cagactcagc 540
cagtgaggag gacttgcala cactggatcg ggagagggct acatttgaga agttgagccg 600
acccccaggc ttttaccctc atcgagttct tttccagaat gccagggcc agtttcttta 660
tgctaccgc tgtgtcctag gccctcatca ggatccccca gaagaggcag aactgctgct 720
acagaacctg caaagtagag gtcccagaga ctgcgtgggtg ctcatggctg cagctgggca 780
ctttgctggg gctatatttc aaggaagaga agtggtgaca cacaaaactt ttcaccgcta 840
tacggttcgg gccaaagcggg gcacagccca ggggcttcgg gatgcccag gtgggccatc 900
acactctgct ggagccaacc tgaggcgcta caatgaagcc acactatata aggatgttcg 960

```

tgacctgctg	gcagggccaa	gctgggctaa	ggcgctggag	gaggctggta	caatactgtt	1020
gcgtgctccc	cgctctggcc	ggtctttgtt	ctttggaggc	aaggagcac	ccctgcaaag	1080
gggggatccc	cgactttggg	atatccccct	cgccaccgc	agaccacct	tccaagagct	1140
acagcgtgtg	ctccataagc	tgaccacttt	gcatgtctat	gaagaagacc	ctcgggaagc	1200
agtcagactg	cactcacctc	agacacactg	gaaaacagta	agagaggaga	gaaagaagcc	1260
tactgaggaa	gaaataagaa	agatctgcag	ggatgaaaag	gaagcgctgg	ggcagaatga	1320
ggaatctccc	aaacaggggt	tgattactat	ctggcaactg	tcagatctga	gtttctgtcc	1380
taaaaatgca	ctggcaaatt	ccctactctc	ataactgacc	catttcctta	tttaaacaca	1440
caccacagt	gtctatctta	gcaattactg	agacatgttt	agtgaccttt	ccaaatctat	1500
tatctctttt	gttatggatg	tttccccaga	gctaagaaaa	tctttgtgtg	gagcagatgc	1560
gtggatgtgt	tgcattggatc	tcacatagtt	tatctttgat	tggatgctgt	ggtttgaggc	1620
tgagggctgc	ttccatgctt	tcctctattg	tgcggcaccc	taaggctact	catacggaga	1680
gtccagggac	ttgttcttgt	ctgggctgag	agcctcatth	ctgcatgtga	tataatggac	1740
tcagctcttg	atgtgctact	tttatctggg	tgacaaaata	aaaacaattt	catttcaggc	1800
tttcattctt	atgcatctgc	ttctaccctc	tgagtactcc	ctagatggat	ttcacatgtc	1860
aggtttttcc	ctaggttcag	ggtcggaggg	agaagatggc	tttcaggtag	agttggagct	1920
agtggagttg	actgtgggga	ctctggatct	ttgtgagtct	gaagtattgc	ccaagcggag	1980
gaggagaaaa	aggaataaga	aggagaaaag	ccgagaccag	gaggctgggg	cacatcggac	2040
tcttctccag	caaactcaag	aagaggagcc	ttccacacag	tcatcccagg	cagttgctgc	2100
ccccttgggc	cctttgctgg	atgaggccaa	agcccctggt	cagccagagc	tctggaatgc	2160
actgcttgct	gcttgccgag	ctggagatgt	tggagtgtcta	aagctgcagc	tagctcccag	2220
ccctgcagac	cctagagttc	tgtctctgct	cagtgcctcc	ttgggctccg	gtggctttac	2280
tctctgcat	gcagcagctg	cagctggaag	aggctcagtg	gttcgtctgc	tgctggaagc	2340
agggtgccag	accattgaca	ccagaaatgg	aggcacggca	ggctacacgg	aaaagggagc	2400
agaaggcagc	ccggcggcaa	cgggaggaac	agcagcagag	gcagcaggag	caggaggagc	2460
gtgaacgaga	agagcagcgg	cgatttgccg	ccctcagtga	ccgagagaag	agagctctgg	2520
ctgcagagcg	ccgactcgct	gccagttgg	gagcccctac	ctctccaatc	cctgactctg	2580
caatcgtcaa	tactcgacgc	tgctggagtt	gtggggcatc	cctccaaggc	ctgactccct	2640
ttcactacct	cgacttctct	ttctgctcca	cacgttgctc	ccaggatcat	cgccgtcagg	2700
caggggaggcc	ctcttctga	tctcttacag	ctctacctgg	ggccaactca	gggacctgag	2760

67

```

agggcacatt cacagcagcc ctaggttttt tcttccccgt gaaaccagag atgatttgga 2820
agatgggggt gaaggacact cgggaactag ggcaaagaca gggctagagg tatgtggagc 2880
tggtactgtc tctggaattt taatcacaat aaagtttggc aaggaatgtg tacttgtact 2940
tacattcaga ggcactgtgg cctttagttc cttaggggtca ccgtggccag caccaaggga 3000
tcctgcccac tccatgtccc aggtccaagg gttacttttt tttttttttt ttttttttga 3060
gatggagtgt ctcaactgtc gccaggctgg agtgcagtgg tacaacctcc actcaactgc 3120
acctctgcct cctggattca agcaattctc ctgccctcag cctcccaagt agctgggatt 3180
agaggtgccc accatcacgc ccggctaatt tttgtatttt tttgagatgg agtgtctcac 3240
tgtcagccag gctggagtgc agtggcacia tctcggtcca ctgcaacctc tgccctccag 3300
attcaagcga ttctcc 3316

```

```

<210> 38
<211> 4629
<212> DNA
<213> Homo sapien

```

```

<400> 38
gcgaattggg cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgtcca gccatgtcgc cggctccaga tgcagccccg gctcctgcgt cgatctccct 180
gtttgacctc agcgcggatg ctccgggtctt tcagggcctg agcctgggtga gccacgcgcc 240
tggggagggt ctggcccggt ctccgcgtac ttctgttcca ggctcagggg agagagaaaag 300
cccagaaaaga aagctactcc agggctctat ggatatttca gagaagttat tttgttcaac 360
ttgtgaccag accttccaga accaccaaga acagagggaa cattataagc ttgactggca 420
tcgggtttaac ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga 480
aaagcagagc tccacagggtg atgagtggta gggggaccta tgtaggagta ggacatggag 540
gtataagaaa gcactagttt agagtttgag gaaagaacaa aaatggggta ccaaacttgt 600
gtatctcata ttgtctgtct tcaggagatc tttccagcat ctcgggatca gaagactcag 660
actcagccag tgaggaggac ttgcagacac tggatcggga gagggctaca tttgagaagt 720
tgagccgacc cccaggcttt taccctcatc gagttctttt ccagaatgcc cagggccagt 780
ttctttatgc ctaccgctgt gtccatggcc ctcatcagag gcaagtgaca gtacagggtg 840
catggctaac cccagccttt tgtacacca gcctagattt cccttagcct agtgcatttc 900
tccctgcttc tcttccttcc tgttggttgt gggatcacg cttgacaaca tagtcttgaa 960
ttgaaactgg cctgaattgt ggaagggtta agtgtaatgg aattaagttt ggaacaaaac 1020

```

68

caatttaggt	ttgaattgtg	attctaccac	ttcatcttga	ctgcctcgcc	aagcctcaaa	1080
tttccttggt	tgtcaaatgg	aaacaagcct	caggggtgga	atattaaatg	agataatggt	1140
tgtgaaaatt	ggcccttagt	ttcttgctct	cagtgaatgc	aacagatatg	tttctgatgc	1200
ttctctgtca	tcaaggatcc	cccagaagag	gcagaactgc	tgctacagaa	cctgcaaagt	1260
agaggtocca	gagactgcgt	ggtgctcatg	gctgcagctg	ggcactttgc	tgggtctata	1320
tttcaagggg	gagaggggtg	tgcatggggg	taaggatgga	gtggatcctg	ttagccaggg	1380
agcaccagct	ggtctccagt	actgagtctg	tgctgtctac	agaagagaag	tggtgacaca	1440
caaaactttt	caccgctata	cggttcgggc	caagcggggc	acagcccagg	ggcttcggga	1500
tgcccagagg	gggccatcac	actctgctgg	agccaacctg	aggcgctaca	atgaagccac	1560
actatataag	gtgagttaag	ccttttagat	ctgggtggtac	tgatccatac	tttttttgca	1620
tctctgttca	actcactggt	aagtagggag	gtttttgact	catatccttc	caattttcat	1680
tataggtaga	caaaggtaga	tgagggattc	cccaggagcc	ttagtcctct	attctttctc	1740
cgtgttatcc	tctactttcc	accccttgca	cttctgtggg	ccattttttt	tagttggaga	1800
tttataaata	gcaagtgaag	gtatgcagga	ctgagttact	ctcttctcta	ggatgttcgt	1860
gacctgctgg	cagggccaag	ctgggctaag	gcgctggagg	aggctggtac	aatactgttg	1920
cgtgctcccc	gctctggccg	gtctttgttc	tttgagggca	agggagcacc	cctgcaaagg	1980
ggggatcccc	gactttggga	tatccccctc	gccaccogca	gaccacacct	ccaagagcta	2040
cagcgtgtgc	tccataagct	gaccactttg	catgtctatg	aagaagaccc	tcgggaagca	2100
gtcagactgc	actcacctca	gacacactgg	aaaacagtaa	gagaggagag	aaagaagcct	2160
actgaggaag	aaataagaaa	gatctgcagg	gatgaaaagg	aagcgctggg	gcagaatgag	2220
gaatctccca	aacaggggtt	gattactatc	tggcaactgt	cagatctgag	tttctgtcct	2280
aaaaatgcac	tggcaaattc	cctactctca	taactgaccc	atttccttat	ttaaacacac	2340
accacagtg	tctatcttag	caattactga	gacatgttta	gtgacctttc	caaactctatt	2400
atctcttttg	ttatggatgt	ttccccagag	ctaagaaaat	ctttgtgtgg	agcagatgcg	2460
tggatgtggt	gcatggatct	cacatagttt	atctttgatt	ggatgctgtg	gtttgaggct	2520
gagggctgct	tccatgcttt	cctctattgt	gcggcacccct	aaggctactc	atacggagag	2580
tccagggact	tgttcttgtc	tgggctgaga	gcctcatttc	tgcatgtgat	ataatggact	2640
cagctcttga	tgtgctactt	ttatctgggt	gacaaaataa	aaacaatttc	atttcaggct	2700
ttcattctta	tgcatctgct	tctaccctct	gagtactccc	tagatggatt	tcacatgtca	2760
ggtttttccc	taggttcagg	gtcggagggg	gaagatggct	ttcaggtaga	gttggagcta	2820
gtggagttga	ctgtggggac	tctggatctt	tgtgagtctg	aagtattgcc	caagcggagg	2880

aggagaaaaa ggaataagaa ggagaaaagc cgagaccagg aggctggggc acatcggact	2940
cttctccagc aaactcaaga agaggagcct tccacacagt catcccaggc agttgctgcc	3000
cccttggggc ctttgctgga tgaggccaaa gcccctgggc agccagagct ctggaatgca	3060
ctgcttgctg cttgccgagc tggagatgtt ggagtgctaa agctgcagct agtcccagc	3120
cctgcagacc ctagagttct gtctctgtc agtgccccct tgggctccgg tggctttact	3180
ctcctgcatg cagcagctgc agctggaaga ggctcagtgg ttcgtctgct gctggaagca	3240
ggtgctgacc ccaactgtgca gtgagtaaag gtccccatcc tgagtcattt tgggtataga	3300
gaggataatt tggagggttaa agttggcatt ggctacttgg caatattccc ttggtctggt	3360
tgggtgatgt tgggaaaggc aggaggcaag ttaaaccat ttttggagtt tttgcaccta	3420
gggactctcg ggcccgcca ccttatactg ttgcggctga caaatcaaca cgtaatgagt	3480
tccgaagggt catggagaag aatccagatg cctacgatta caacaaggct caggctcatct	3540
ggaataggaa ggacaagcag agtgggaagg catcacagat cctggctaga tcccttctac	3600
ttcttgcatg tttaccaaaag ataacttata gggctctata tttgaaatcc tccaggtgcc	3660
aggaccattg acaccagaaa tggaggcacg gcaggctaca cggaaaaggg agcagaaggc	3720
agcccggcgg caacgggagg aacagcagca gaggcagcag gagcaggagg agcgtgaacg	3780
agaagagcag cggcgatttg ccgccctcag tgaccgagag aagagagctc tggtctcaga	3840
gcgccgactc gctgccagt tgggagcccc tacctctcca atccctgact ctgcaatcgt	3900
caataactcga cgctgctgga gttgtggggc atccctccaa ggctgactc cctttcacta	3960
cctcgacttc tctttctgct ccacacgttg cctccaggat catcgccgtc aggcaggggag	4020
gccctcttcc tgatctctta cagctctacc tggggccaac tcagggacct gagagggcac	4080
attcacagca gccctagggt ttttcttccc cgtgaaacca gagatgattt ggaagatggg	4140
ggtgaaggac actcggaac tagggcaaag acagggctag aggtatgtgg agctgggtact	4200
gtctctggaa ttttaatcac aataaagttt ggcaaggaat gtgtacttgt acttacattc	4260
agaggcactg tggcctttag ttccttaggg tcaccgtggc cagcaccaag ggatcctgcc	4320
cactccatgt cccagggtcca agggttactt tttttttttt tttttttttt tgagatggag	4380
tgtctcactg tcagccaggc tggagtgcag tggtaacaac tccactcact gcaacctctg	4440
cctcctggat tcaagcaatt ctctgccct cagcctccca agtagctggg attagaggtg	4500
cccaccatca cgcccggtta atttttgtat ttttttga ga tggagtgtct cactgtcagc	4560
caggctggag tgcagtggca caatctcggc tcaactgcaac ctctgcctcc gagattcaag	4620
cgattctcc	4629

<210> 39
<211> 4090
<212> DNA
<213> Homo sapien

<400> 39
gcgaattggg cctctagacg catgctctag caggcggcgc agtgtgctgg aaaggaccaa 60
tcgacctgag gacttgcgag ccgctcagct cccgggacgt ttggagctgc tgctaaataa 120
tttctgctca gccatgtcgc cggctccaga tgcagccccc gctcctgcgt cgatctccct 180
gtttgacctc agcgcggatg ctccggtctt tcagggcctg agcctggtga gccacgcgcc 240
tggggaggct ctggcccggg ctccgcgtac ttctgttca ggctcagggg agagagaaag 300
cccagaaaga aagctactcc agggctcctat ggatatttca gagaagttat tttgttcaac 360
ttgtgaccag accttcaga accaccaaga acagagggaa cattataagc ttgactggca 420
tcggtttaac ctaaagcaac gtctcaagga caagcctctc ctgtctgccc tggactttga 480
aaagcagagc tccacaggag atctttccag catctcggga tcagaagact cagactcagc 540
cagtgaggag gacttgca cactggatcg ggagagggct acatttgaga agttgagccg 600
acccccaggc ttttaccctc atcgagttct tttccagaat gcccagggcc agtttcttta 660
tgctaccgc tgtgtcctag gccctcatca ggatcccca gaagaggcag aactgctgct 720
acagaacctg caaagtagag gtcccagaga ctgcgtggtg ctcatggctg cagctgggca 780
ctttgctggt gctatatttc aaggaagaga agtggtgaca cacaaaactt ttcaccgcta 840
tacggttcgg gccaaagcgg gcacagccca ggggcttcgg gatgcccag gtgggcatc 900
acactctgct ggagccaacc tgaggcgcta caatgaagcc acactatata aggatgttcg 960
tgacctgctg gcagggccaa gctgggctaa ggcgctggag gaggctggtg caatactggt 1020
gcgtgctccc cgctctggcc ggtctttgtt ctttgaggc aaggagcac ccctgcaaag 1080
gggggatccc cgactttggg atatccccct cgccaccgc agaccacct tccaagagct 1140
acagcgtgtg ctccataagc tgaccacttt gcatgtctat gaagaagacc ctcggaagc 1200
agtcagactg cactcacctc agacacactg gaaaacagta agagaggaga gaaagaagcc 1260
tactgaggaa gaaataagaa agatctgcag ggatgaaaag gaagcgtg ggcagaatga 1320
ggaatctccc aaacagggtt tgattactat ctggcaactg tcagatctga gtttctgtcc 1380
taaaaatgca ctggcaaatt ccctactctc ataactgacc catttcctta tttaaacaca 1440
caccacagt gtctatctta gcaattactg agacatgttt agtgacctt ccaaatctat 1500
tatctctttt gttatggatg tttccccaga gctaagaaaa tctttgtgtg gagcagatgc 1560
gtggatgtgt tgcattggtc tcacatagtt tatctttgat tggatgctgt ggtttgaggc 1620

71

tgagggctgc ttccatgctt tcctctattg tgcgggcaccc taaggctact catacggaga	1680
gtccagggac ttgttcttgt ctgggctgag agcctcattt ctgcatgtga tataatggac	1740
tcagctcttg atgtgctact tttatctggg tgacaaaata aaaacaattt catttcaggc	1800
tttcattctt atgcatctgc ttctaccctc tgagtactcc ctagatggat ttcacatgtc	1860
aggtttttcc ctaggttcag ggtcggaggg agaagatggc tttcaggtag agttggagct	1920
agtggagtgt actgtgggga ctctggatct ttgtgagtct gaagtattgc ccaagcggag	1980
gaggagaaaa aggaataaga aggagaaaag ccgagaccag gaggctgggg cacatcggac	2040
tcttctccag caaactcaag aagaggagcc ttccacacag tcatcccagg cagttgctgc	2100
ccccttgggc cctttgctgg atgaggccaa agcccctggt cagccagagc tctggaatgc	2160
actgcttgct gcttgccgag ctggagatgt tggagtgtta aagctgcagc tagctcccag	2220
ccctgcagac cctagagttc tgtctctgct cagtgtcccc ttgggctccg gtggctttac	2280
tctcctgcat gcagcagctg cagctggaag aggtcagtg gttcgtctgc tgctggaagc	2340
aggtgctgac cccactgtgc agtgagtaaa ggtccccatc ctgagtcatt ttgggtatag	2400
agaggataat ttggagggtta aagttggcat tggctacttg gcaatattcc cttggtctgg	2460
ttgggtgatg ttgggaaagg caggaggcaa gttaaaccga tttttggagt ttttgcacct	2520
agggactctc gggcccgcc accttatact gttgcggctg acaaataaac acgtaatgag	2580
ttccgaaggc tcatggagaa gaatccagat gcctacgatt acaacaaggc tcaggtcac	2640
tggaatagga aggacaagca gagtgggaag gcatcacaga tcctggctag atcctttcta	2700
cttcttgtag ttttaccaaa gataacttat agggctctat atttgaaatc ctccaggtgc	2760
caggaccatt gacaccagaa atggaggcac ggcaggctac acggaaaagg gagcagaagg	2820
cagcccgccg gcaacgggag gaacagcagc agaggcagca ggagcaggag gagcgtgaac	2880
gagaagagca gcggcgattt gccgccctca gtgaccgaga gaaggtgagg ctggagggttc	2940
tcttgtccat ggcaaggctt cctagaggtc taacccttc cccagcgtgc agctgtcacc	3000
tcttgggtgc cctactcgcc actggggcct tgtccttaac acaacttgc tccctcagag	3060
agctctggct gcagagcgcc gactcgtctc ccagttggga gccctacct ctccaatccc	3120
tgactctgca atcgtcaata ctcggtatgg ggtgcgggat gaggagtggt gtggactgta	3180
atgttgacag gaccattggg ggcaagagaa atgagtacct gcacagtatt cagaagggtg	3240
gttgagggaa tttgggatgt tctggcagat gctgcagtga aaacagaaat ggcaggagga	3300
gggacgggga gagcgtggat tggaaagttt ctgggggtacc tgtccagcca ttttcttcc	3360
tcttgttcag acgctgctgg agttgtggg catccctcca aggcctgact ccctttcact	3420
acctcgactt ctctttctgc tccacacgtt gcctccagga tcatcgccgt caggcaggga	3480

72

```

ggccctcttc ctgatctctt acagctctac ctggggccaa ctcagggacc tgagagggca 3540
cattcacagc agccctaggt tttttcttcc ccgtgaaacc agagatgatt tggaagatgg 3600
gggtgaagga cactcgggaa ctagggcaaa gacagggcta gaggtatgtg gagctggtag 3660
tgtctctgga attttaatca caataaagtt tggcaaggaa tgtgtacttg tacttacatt 3720
cagaggcact gtggccttta gttccttagg gtcaccgtgg ccagcaccaa gggatcctgc 3780
ccactccatg tcccaggtcc aagggttact tttttttttt tttttttttt ttgagatgga 3840
gtgtctcact gtcagccagg ctggagtgcg gtggtacaac ctccactcac tgcaacctct 3900
gcctcctgga ttcaagcaat tctcctgccc tcagcctccc aagtagctgg gattagaggt 3960
gccaccatc acgcccggct aatttttgta tttttttgag atggagtgtc tcaactgtcag 4020
ccaggctgga gtgcagtggc acaatctcgg ctcaactgaa cctctgcctc cgagattcaa 4080
gcgattctcc 4090

```

```

<210> 40
<211> 3770
<212> DNA
<213> Homo sapien

```

```

<400> 40
ctacttgccg ggctgagatg ggtggatcgc ttgaacacag gaggttaagg ctgcagtgag 60
ccgagatcgt gccactgcac actagcctgg gcggcgggag tgagaccgtg tctcaaaaac 120
aacaacaaca acaacaacaa caaaaaacaa aaaatggacc gaggttgggg tggcaaggct 180
gggctggaat ggaggctgag gacaggccct ggcagcgcgc agggggtaga gggcgtgcag 240
caggaagcaa gggccaccct cagggatggg ggactgggccc tggcagctgc cgaggggtag 300
gcaggatgca tgtggagagc cacgtggggc tgcggtgag ccaaggattg tgagaggaat 360
cctgctgctc tgggtggcag gcagaaggcc tggggccacag gtggggcctg ctggggagac 420
cacaggtagg gcaggtatgg ctgaccctgg gcgctcggca caccggtact cctcctgcct 480
ggctgctccc tcagcagggtg tcagagtctg agtcaccttt tcctcacgtg acccccacca 540
ggaccaggca gggagggtggc catctttggc cctgactcag ccctgcaggg agaagagggg 600
ctggagggtg aggtggcctg ccactgtctc ccccaacccc tgcccacctg ggtcttgtgc 660
tgtgtgtact tctgctgctg ctgtcctccc ttgtgaccag cttctgctgg acctgattgt 720
ccagccatga tcttctgatc atcatgggtt ctggacacac atctgttgga ctaaggaatt 780
tgtgttctct gccacagaac agtgccacct ctgcctcaag cagttcctgg tctgggcact 840
ccccaggga ggccggagag gactgagccg tctcctgatg cagaaagtca gcccgcctgg 900
gcagtgtctg cccaggaagt agggccccac ccaacgggtg gctgactgtg gaacactgtg 960

```

tctgcctagg	cgggtctggc	cacattcctg	cagcatgatg	gttggcagtt	agcttcctgg	1020
agtgtggacg	ggtccttccc	tgcctgactg	acttgtccac	agggcagcaa	gaagatgcgc	1080
cccttggtgt	gtcccgagtg	tctcctgtag	gcagctgagc	tttcgcaggt	ctggctggga	1140
gccaaggtta	ccaatctggg	gggtaaccca	gccatcatct	ggatgagcgt	tggggctctt	1200
ctgctctgga	ttgcagggca	gacgtcgttt	tgtgctggtc	aggcccctcg	cctgcctgat	1260
tcccttgagg	ccccccaaga	caggagatgg	gcagatgtcc	tctgaagtgg	catgggctgc	1320
ttgctcactg	gccttcccag	gaccctgccc	agggtgtgct	gcttggcccc	aggtcgaatt	1380
cctgtgctgg	cagcaagcag	gggcctgggc	cgtcggctgg	ctggggctca	tgcagccatc	1440
ccctttgcag	ccatccgggt	gacctgcata	ggttcctgcg	gggtctccaa	caaggcta	1500
gacacagcgt	gggtagtggg	ggaggggttac	ttcaacagtt	ccctgtcgct	ggcagacaag	1560
ggtaagtttg	ggagccagtt	cccagatggg	gaccctggg	ggcagccctt	ggagtggggc	1620
ctgtctgtcc	tgtcctcccc	ctttcctaga	tctgcctctt	ggatctggca	ctggcccctc	1680
ttgtcccagg	gttgtggggc	ctgatggctt	tctcataggc	gctgggtggc	tgaagggggc	1740
acctccccct	tgtgctggat	ttgggtgggc	tctgaccctt	ggcacagctg	tttcaccctt	1800
gccccattcc	caaaccactc	tctgttcaa	gccaagctct	tatcctcagt	ccttcaccca	1860
agccacagcc	ctgtgctgct	gtggcgctgg	tgactgagac	agggagaccc	agtgtgagct	1920
ggtcagctgc	agtgccact	cttatcctgg	ccctgtggcc	acctcaggcc	acctcagggg	1980
gggactgccc	cccacctccc	tctgtctct	gcagggagcc	tgcccgctgg	agagcacagc	2040
ttccccctcc	agttcctgct	tctgggtgag	agcccagcct	ggacaggcct	ggtgatgacc	2100
aactggctct	gggaggtggg	ctctgcaggg	tgggaaggcag	ccaggtacca	gtgcctgctc	2160
tctccccagc	cactgcaccc	acgtcctttg	agggtccttt	cgggaagatc	gtgcaccagg	2220
tgagggccgc	catccacacg	ccacggtttt	ccaaggatca	caagtgcagc	ctcgtgttct	2280
atatcttgag	ccccacttc	ttggaccccg	tcttcctctc	caccaagagc	cattcgcagc	2340
ggcagccctt	gctggccacc	ttgagttctg	tgcctgggtg	gccggagccc	tgcctcagg	2400
atggcagccc	tgcctcacac	ccgtgcacc	ctcccttgtg	catttcaaca	ggtgccactg	2460
tcccctactt	tgcagagggc	tccggggggc	cagtgcacac	taccagcacc	ttgattcttc	2520
ctccagagta	cagttcttgg	ggctaccctt	atggtgagtc	gacagccagg	gcttggcagg	2580
gaggggacgc	caagagcccc	acgcagaccc	tgctttcttc	ccgcagaggg	cccaccgtct	2640
tatgagcaga	gctgcggcgg	cgtggaaccc	agcctgaccc	ctgagagctg	accccgctct	2700
gccttctcca	ggcaggcctg	gcctctgccc	tgggactggg	gcgcccaggg	cctcgtgcct	2760

74

tctctcttgg cctagcctgg ccactcagg acctgcccag cctctgccag ctctcttggc 2820
 atccgccctc ttctccctgg ggctggggtg ggggtggcag ggagctggga cctggagaga 2880
 caactcctgt aaataaaaca ctttatttgt agagctgggc ctaccggcc tcgccttggg 2940
 caccctccag cctctggcag ggcctgacct tgaccccatc ctgggctctg ctcatgccag 3000
 ggcagacctt gctcctcagg tgccccctac taccctgggg gcgcgctcag gctaccatgc 3060
 ctgggcggag cgcgggggct ggctcccttt aacagaggac gggggctggc gctgggcctt 3120
 ccagagccgc ctaaggacac gtccttggag cccggctcat gaaatggggc ctacgcaac 3180
 cttggtgttt ggttaaattt gtgccttggc tcctgttgcg tttaggcacc caggttccca 3240
 agcgggtgtg tgtggacctc ccctgctgag aaaccgctg gcgcgtggag gggcttgccc 3300
 aggtccctcc gacaggagac cctgctccct gggccttggc ccctgctagc accactgcca 3360
 ggcctggctc tgagggacgc tgcttctccc ctgccacgt ctccaaggcc tgaaaggcag 3420
 ggttctgttg aacagcattt ctcaaagggg catcctgcag acgttccagc ccagggggag 3480
 gccccgtgat gctggctggc ctgcctggag ggtgcttcca gaaacaacgg agatgggtgt 3540
 gaagaaggct ctgggttcaa attccagttc tcctttaggg cgaaggggtg actgtggcta 3600
 ggtttgtcag ttaccttacc tataaaatgc acaggaccag gccgggcacg gtggctccct 3660
 cctgtaatcc cagtactttg ggaggctgag gcagggtgat tgcttgaggt caggagtccc 3720
 agatcagcct gggcaacata gtgaaactcc atctgtacaa aaaaaaaaaa 3770

<210> 41
 <211> 1650
 <212> DNA
 <213> Homo sapien

<400> 41
 agggaccgta tttcaagtcc tgccctccca ccagcgctg cggcccgag cagtgtccgg 60
 ctccctgccg acgccctgga gctgggacgg cctggggcgc tggtcggccg ccttccccca 120
 gaaggaggcg tgagcggggc gtggccggaa cgtctctctt cggagccgcg ccccgggcgg 180
 gcacagttac gcgccgcacg gatcatggcc gcagcgctc tggggcaggt gggttcgag 240
 ccccgcgccg gggctctgga aataaccggg ggaggcgac ttccgggaag accactggcc 300
 tgcagccggt ccgactcggc gccgctcctc ccgcctggg gttccgctgc agggcgcgga 360
 ggaggatggc gccctttgga aaggagcacg aggccttaat ggggtgcagc tgtttgaggg 420
 catgaaggcg ttcaaaggca aagaccagca ggtgcgcctc ttccgcccct ggctcaacat 480
 ggaccggatg ctgcgctcag ccatgcgcct gtgcctgccg agtttcgaca agctggagtt 540
 gctggagtgc atccgcggc tcacgaagt ggacaaggac tgggtccccg atgccggcg 600

75

caccagcctc tatgtgcggc ctgtgctcat tgggaacgag ccctcgctgg gtgtcagcca 660
 gccacgcgc gcgtcctgt tegtattct ctgccagt ggtgcctact tccctggagg 720
 ctccgtgacc cgggtctccc tectggccga ccagccttc atccgggcct ggggtggcgg 780
 ggtcggcaac tacaagttag gtgggaatta tgggccacc gtgttagtgc aacaggaggc 840
 actcaagcgg ggctgtgaac aggtcctctg gctgtatggg ccgaccacc agctcaccga 900
 ggtgggaacc atgaacatct ttgtctactg gaccacgaa gatgggggtgc tggagctggg 960
 gacgcccccg ctgaatgggt ttatcctgcc tggagtgggc agacagagtc tactggacat 1020
 ggctcagacc tggggtgagt tccgggtggg ggagcgcacg atcaccatga agcagttgct 1080
 gcgggcccctg gaggagggcc gcgtgcggga agtctttggc tcgggcaccg cttgccaggg 1140
 ctgccagt gaccgaatcc tgtacaaaga caggaaacct cacattcca ccatggaaaa 1200
 tgggcctgag ctgatcctcc gcttcagaa ggagctgaag gagatccagt acggaatcag 1260
 agccacgag tggatgttcc cgggtgtgaag ctgcaggctg tgctccagat ccaccgacct 1320
 gtagcatctc gtcacgccag cactgcctc cctaccaatg actcacctga agtgaatac 1380
 gaaataaaag gccagcgggc ggctctggg tctctggcgc ccccatgtgg ttgcgacact 1440
 cccaaagccg taagggccga ccaggcatc ttggcccca gccctcgtc gcgggttcag 1500
 gtccgcccac tactccttgt cgtgcggtca aggatacacc ttggccccga ttccggatct 1560
 ctccgttctc aggccagacc cctggtgctg ccgttgattt ttttttctct gtctttgctg 1620
 caattttgaa ataaaatgcc aaagaacaca 1650

<210> 42
 <211> 2040
 <212> DNA
 <213> Homo sapien

<400> 42
 tccaaattgt ttatttat atacaaagaa ttatcagggt tttccatgga gtaaatgccc 60
 cggataatcc tctgaaggaa gagcatttag tccaacttaa tgaaaccgat atccttcgcg 120
 tactggacgg aaacactggc ggcacatatg gaggccatat tcccggatca gaccgtgccg 180
 gtttgaacag acacgacaag agcgagaacc cgggccgaat ttacgcgggt ggctccagta 240
 cagctgctgg tgacctct ttgctctcag ccgcacggat catggccgca gccgtctggt 300
 ggcagatctg ggcacgaaag cttctctctg tcccttggct tctgtgtggg ccagaagat 360
 atgcctctc cagtttcaag gctgcagacc tgcagctgga aatgacacag aagcctcata 420
 agaagcctgg ccccggcgag cccctggtgt ttgggaagac atttaccgac cacatgctga 480
 tgggtggaatg gaatgacaag ggctggggcc agccccgaat ccagcccttc cagaacctca 540

76

cgctgcaccc agcctcctcc agcctccaact actccctgca gctgtttgag ggcataaagg 600
 cgttcaaagg caaagaccag cagggtgcgc tcttccgccc ctggctcaac atggaccgga 660
 tgctgcgctc agccatgcgc ctgtgcctgc cgagtttoga caagctggag ttgctggagt 720
 gcatccgccc gctcatcgaa gtggacaagg actgggtccc cgatgccgcc ggcaccagcc 780
 tctatgtgcg gcctgtgctc attgggaacg agccctcgct ggggtgcagc cagcccacgc 840
 gcgcgctcct gttcgtcatt ctctgccag tgggtgccta cttccctgga ggctccgtga 900
 ccccggtctc cctcctggcc gaccagcct tcatccgggc ctgggtgggc ggggtcggca 960
 actacaagtt aggtgggaat tatgggccc cctgttagt gcaacaggag gcaactaagc 1020
 ggggctgtga acaggctcctc tggctgtatg ggcccgacca ccagctcacc gaggtgggaa 1080
 ccatgaacat ctttgtctac tggaccacg aagatggggt gctggagctg gtgacgcccc 1140
 cgctgaatgg tggtatcctg cctggagtgg tcagacagag tctactggac atggctcaga 1200
 cctgggtgag gacatggcat cttctggta tgggcgccat gtgccagggg ccagggcac 1260
 agagggtgag gactggggca cactggcatg tctctgcccc ttctcctggg tctgtctctc 1320
 ccgttgggccc tctgtcttct tcaactatcct caggagaggga gaggtggagg tctgcagcag 1380
 ggcaaccctc tggggacacg tgtctctgcc aactgcctg cagggtgagt tccgggtggt 1440
 ggagcgcacg atcaccatga agcagttgct gcgggcccctg gaggagggcc gcgtgcggga 1500
 agtctttggc tcgggcaccg cttgccagggt ctgcccagtg caccgaatcc tgtacaaaga 1560
 caggaacctc cacattccca ccatggaaaa tgggcctgag ctgatcctcc gcttccagaa 1620
 ggagctgaag gagatccagt acggaatcag agcccacgag tggatgttcc cgggtgtgaag 1680
 ctgcaggctg tgctccagat ccaccgaccc gtagcatctc gtcacgccag cactcgcctc 1740
 cctaccaatg actcacctga agtgcaatac gaaataaaaag gccagcgggc ggcgtctggg 1800
 tctctggcgc ccccatgtgg ttgcgacact cccaaagccg taagggccga cccaggcatc 1860
 ttggccccca gcccctcgct gcgggttcag gtccgcccct tactccttgt cgtgcgggtca 1920
 aggatacacc ttggccccga ttccggatct ctccgttctc aggccagacc cctggtgctg 1980
 ccgttgatctt ttttttctct gtctttgctg caattttgaa ataaaatgcc aaagaacaca 2040

<210> 43
 <211> 1191
 <212> DNA
 <213> Homo sapien

<400> 43
 tactcgtgtg ttacagtgtg cccccagatg ttctcccaat tgcgtggcgt tcaactgcagc 60
 tgtaacggtc cctgcgtgca tttatgcgct ctttggaccc tgtgaacaca ccgggatact 120

77

ggtaatcttg	ccccctatgg	agtacttatg	gcgctccct	gtgtttat	wttkggtat	180
aaacccttwa	taascawttt	aacwyacaka	awagttcawa	gawwaagatc	agaaagaata	240
ctgcttcata	tgtcccaaaa	gagaaaaaaa	taaaggggac	aatgccaaca	tgctcaacaa	300
taaaggcttc	tttttcttat	ttttttaata	caaaatacaa	gcaaaggata	cacatactta	360
aaacagagct	caggagcaga	cacgcagtcc	tggaaaccct	tcaataaaag	caaagcagga	420
gtttgttttt	tctttgtcta	tgcagatata	tacagagact	gggatatgta	aaaattaagt	480
atcacaaaag	accatcacac	gattctacca	atgcatgttg	catctgtaat	tcacgaacat	540
ggtcaacaaa	atcatgttca	cttcaacccc	atttcattta	aattaaaga	aaaaaccttt	600
taaataaagt	ggttacattc	aaactttaac	ttccttagta	ccatgctgca	gatttcagca	660
ctgttaaggt	attgcaagaa	tgcccaaccc	tctggtgtct	gatcatgtat	ctagcaacat	720
tgcagtatga	agaaaagaga	tgccccggtc	tcagcccatg	gactagttaa	tacagtgaag	780
caggttcctg	tcttttacc	ttcctgctca	gaacataaaa	gattaaggac	taaaatcaag	840
gaagactggg	agtttttagag	ctggcaaaat	gaagtctaaa	agataaatca	aggcaaacaa	900
ttactgagaa	cttggctgtt	gcttaacctg	gcaagtctaa	aagcctttct	ttaaccttgt	960
aggaattaga	tgcataaggt	ttgctgcaac	atgttcatgg	taaacaaact	aagtagagct	1020
cttatttaca	aatcttgtaa	caaatacttc	tggaggaaaa	agagaaaaga	attcactaag	1080
ttccagaaga	caagcttaat	tgcagacgta	tacaaacaca	catcacacgt	aacaccacac	1140
atattcgggt	ttttatacat	gttatttttag	ggcataagct	gagtactata	c	1191

<210> 44

<211> 3097

<212> DNA

<213> Homo sapien

<400> 44

gagcctacgt	gacacgacgt	gaagtggcaa	tggcggcgga	cgagggtgt	gccccggcc	60
cgccccgtta	ggagccgcga	gtaaggcgag	tgtccaacgc	cgagctggct	gacagagccc	120
ggccgcggcc	tccgcgggcc	caggggcctc	caggccccgt	gaccaccgga	ccgagcaccc	180
tggaacggcc	ccagcttggc	ctcgggacag	tccgcgcgct	caccgactcc	ctggtgaacg	240
ccgcttggcc	gccgccgccg	cgcaggacc	cacgagaagc	ggaaaccggc	gcaaggaccc	300
ggtccccccg	gcgcctacg	tggagtggag	ctgcggctcc	gccccgcg	ctccgattgg	360
ctctggggcc	tggcccgccc	cttcccgaca	ctgtcattgg	cctaggcaaa	gccgtaaggg	420
tggggaatcc	tattggccca	ggggtgccgc	ttgtcctctc	ccggtgctct	cattggccga	480
gcgcggccgt	tgggggtgag	gccgcgtcgg	gggaggacaa	caaagggccg	cgggcggcgg	540

gcagtgggtgt cccagtctcc cggtgcttcc ctgaggctga ggcgcccggc ctcccggccg	600
ccgcgctcca gatgaagtgt gagcactgca cgcgcaagga atgtagtaag aaaacaaaaa	660
ctgatgacca agagaatgtg tcagccgatg caccgagtcc agcccaggaa aatggagaga	720
agggagaatt ccacaagttg gctgatgcca agatattttt gagcgactgc ctggcatgtg	780
acagctgtat gactgcagag gaaggagtcc aactttccca gcaaaatgcc aaggacttct	840
tccgcgttct gaaccttaac aagaaatgtg atacctcaaa gcacaaagtg ctggtagtgt	900
ctgtgtgtcc tcaatctttg ccttattttg ctgctaaatt caacctcagt gtaactgatg	960
catccagaag actctgtggt ttctcaaaa gtcttggggg gcactatgta tttgatacga	1020
cgatagctgc ggatttttagt atcctggaga gtcaaaaaga attcgtgcgt cgctatcgcc	1080
agcacagtga ggaggaacgc accctgcccc tgctgacctc tgctgtcct ggctgggtcc	1140
gatacgccga gcgggtgctg ggtcgcccc tcaactgccc cctctgcacc gccaagtccc	1200
cccagcaggt catgggctct ttggtgaagg attatttcgc cagacagcag aacctgtctc	1260
cagagaagat tttccacgtc attgtggccc cttgttatga caagaagctg gaggtctctc	1320
aggaaagcct tccccctgct ttgcatggct cccggggcgc tgactgcgtg ttaacatcag	1380
gtgaaattgc tcaaataatg gagcaaggtg acctctcagt gagagatgct gccgtcgaca	1440
ctcttgtaag tggcttctct ggggagagtc ctctgggggg caggacctct cgtcagccct	1500
gccagccacc tgccaggccc agggctgccc tgctttatga cgaggccatg tgaagcactg	1560
cgatgcagtc cccagggccc ctgccacaac cctcagccac cctatagctg ttgaatggga	1620
taccattcac cagggtgccc agggagtga aggtttggcg ggcttccatg gtgtgattgc	1680
caccctaaac ccttgaggaa ggacaaaatg cacaaacagc tttgtcctcg ctggcgagg	1740
gctgttcatg tgtgcaccag ccccttctt cctgagtgc gtctgtgcct cacattccgt	1800
gtggtctatg gggcagtggg gcagctccct caggccctt tgcaccaaga gtggcagtct	1860
gggttctgct gggggcggtg acctcgatac ccagcaaggc ccagggtcc tgctgcacc	1920
caggcactgg gcagagagca ccacaaaggc agcagcgtgt gtctcgcca acagtccgca	1980
tcccaggtgc tgcattcgtg cccttttggg ggaaggtaag atagaagttt acagaattta	2040
agcatctgtg gggccaggcg tgggtggctca cacctataat cacagtactt tgggaggctg	2100
aggtaggagg atcacttgag tccaggaggc tggggctgta gtgagctatg attgtgccac	2160
tgactccag cctgagagac agaacaagac cctgtctcca aaagaaaaag accataagca	2220
tctggagcac ctgcaagagc ctgtgcatgg gctcggcggtg actgagagtc ctccgccttg	2280
cccagcatc aggcaggccc tcccaagacc cctcaagatg ctagaatggg actttctgca	2340
aagtttaaac atctgtacag gatttagctt ctttgggtggc gaacttggat tgtgatgccc	2400

tcaggaagtc tggtttcact ctttctgtca ctgtacgtca gcattcatga aggacttgta	2460
ttttctctgc cttgtgaggt ttggagactt gaaggaggac aaagtgacgc gtcgatgagg	2520
agccagctca gacgggcacc tggcacacat cttcagacat gcggccaagg agctgttcaa	2580
cgaggatgtg gaggagaaac aaagacttcc aagagggtcac ccttgagaag aacggagagg	2640
tgggtgttacg ctttgcctgca gcctatggct ttcgaaacat ccagaacatg atcctgaagc	2700
ttaagaagggt caagttccca ttccactttg tggagggtcct cgctgtgtgct ggaggatgct	2760
taaatggcag aggccaaagcc cagactccag acggacatgc ggataaggcc ctgctgcggc	2820
agatggaagg catttacgct gacatccctg tgcggcgctcc ggagtccagt gcacacgtgc	2880
aggagctgta ccaggagtgg ctggagggga tcaactcccc caaggcccga gaggtgctgc	2940
ataccacgta ccagagccag gagcgtggca cacacagcct ggacatcaag tggatgaagtc	3000
aggccagggc cttccagctg ctcttggggc cagagccaag agcctctcag tagagggagg	3060
ggctgccctg agtggagtat taaagacact taagaaa	3097

<210> 45
 <211> 3709
 <212> DNA
 <213> Homo sapien

<400> 45	
gagcctacgt gacacgacgt gaagtggcaa tggcggcgga cgagggtgt gcccccgcc	60
cgccccgtta ggagccgcga gtaaggcgag tgtccaacgc cgagctggct gacagagccc	120
ggccgcggcc tccgcgggccc caggggcctc caggccccgt gaccaccgga ccgagcacc	180
tggaaacggcc ccagcttggc ctgcggacag tccgcgcgct caccgactcc ctggtgaacg	240
ccgcttggcc gccgccgccc ccgcaggacc caccgagaagc ggaaaccggc gcaaggaccc	300
ggcccccccg gcgccgtacg tggagtgagc ctgcggctcc gccccgcgcg ctccgattgg	360
ctctggggcc tggcccgccc cttcccgaca ctgtcattgg cctaggcaaa gccgtaagg	420
tggggaatcc tattggccca ggggtgcggc ttgtcctctc ccggtgctct cattggccga	480
gcgcggccgt tgggggtgag gccgcgtcgg gggaggacaa caaagggccg cgggcggcgg	540
gcagtgggtgt ccagctctcc cgggtgcttc ctgaggctga ggcgcccggc ctcccggccg	600
ccgcgtcca gatgaagtgt gagcactgca cgcgcaagga atgtagtaag aaaacaaaa	660
ctgatgacca agagaatgtg tcagccgatg caccgagtcc agcccaggaa aatggagaga	720
agggagaatt ccacaagttg gctgatgcca agatattttt gagcgactgc ctggcatgtg	780
acagctgtat gactgcagag gaaggagtcc aactttccca gcaaaatgcc aaggacttct	840
tccgcgttct gaaccttaac aagaaatgtg atacctcaaa gcacaaagtg ctggtagtgt	900

ctgtgtgtcc tcaatctttg ccttattttg ctgctaaatt caacctcagt gtaactgatg	960
catccagaag actctgtggt ttctcaaaa gtcttgggggt gcactatgta tttgatacga	1020
cgatagctgc ggatttttagt atcctggaga gtcaaaaaga attcgtgcgt cgctatcgcc	1080
agcacagtga ggaggaacgc accctgccc tgcctgacctc tgcctgtcct ggctgggtcc	1140
gatacgccga gcgggtgctg ggtcgcccca tcaactgccc cctctgcacc gccaagtccc	1200
cccagcaggt catgggctct ttggtgaagg attatttcgc cagacagcag gtaagctgac	1260
ctctctggag ggcaggaagg gcaggtgagg gaggaccatg gcacaggggc ccaggaaggg	1320
gaggttcccc atctgtccag cgtggtcacg gccccccagg tgagcaaaaa ggaagctgtg	1380
gcttcaggaa acatcagcag tgctgtttct caccaacgcc cgtctctatt cctcccttcc	1440
tcttacaacc tttttctttc ttcttaacct tttctgtggc atctggtggt tttcctgggc	1500
ccattcgggg tagagtagaa atgctgcca ccctgctgag aatgaggccc cggaccttgc	1560
ccagagatgc tgcctgcctg ctgttgggat gcctcaccca aggaggccta aagtggcctg	1620
tgcctggagg ctgagggcag ctgtcattta gggagccaaa cctggcctct gccctggcct	1680
ttccctcaga ccacgggcat cacatgactt ccttagccac ttctgtgaga ttccacagtg	1740
gccatgatgc tgtcctctc ttgatggagt gagtcccggc ctgaggctcc tgcgtgcccc	1800
ctgcctcggc ccgtcgtgtt ctcccttccc acaacctctt gcggccccca ggctctccca	1860
gaactccctg atgtctccaa ggccaccctc aaaggggacc taatcccatg gtttattctc	1920
cagccgctgg catgccccac actgctcttc tcaatcagac tttacacctc cctccgggga	1980
agcagccatg tccaggcact gtggttaactg gagtagaggg taacatcagg atctttcaga	2040
catttgaagt cagcctaaaa aggcacgcac atttggcagc cttgttattt tagcaatagg	2100
cagtgttttc cttctatcta gtttctttct acccataaag tcataatttg gcaaattata	2160
aacatgatgt cagagtatac acgttcagtt taggcaggac cgaaaaatga gctatgggac	2220
tttctgtccc ttgccctgcc ccacctgaa gagagccatg tgtttctagg tcttcacgt	2280
gctgccctct cccctggatg gagccagag ctatgtgtcc tcctgtaggt gccatgaggg	2340
gcagccgag agacccggcc tccagcagcc tctcacctga ccacctggca cagctgcaac	2400
caggggcggg tttcctcttt tgcctcgccc ctgccctgag tgggtccttt ccaccagac	2460
ctcaccacag ggtcacttgg caactgctgg ctctggctt tgccaccagt tcaggccag	2520
ccacaggcac ctgctctgtg tggaccaggg agaggggtag cttgggtggg catcgcttgg	2580
gaagattcag gaaagtctgt ggggtgaagta gccagcgcgc acacacacac acacacacac	2640
acactcactc aggcctcctg caatggctcc atgctggacc agcaacagct ccacgcccgc	2700

81

tggacttcct cttacttggt ttgtcccat gtgcacaggt tcaggaggaa atgggctcag	2760
ccttcccggg tcttccagga ggctccactg gctggggggc tttgcagatg tttgggagga	2820
ccacaccag tcctgaggga ggcagcagac agacatggat ggagtgtggt tgctcgaacc	2880
tgtctccaga gaagattttc cacgtcattg tggccccttg ttatgacaag aagctggagg	2940
ctcttcagga aagccttccc cctgctttgc atggctcccg gggcgctgac tgcgtgttaa	3000
catcagggtga aattgctcaa ataatggagc aagggtgacct ctcaagtgaga gatgctgccg	3060
tcgacactct gtttgagagac ttgaaggagg acaaagtac gcgtcatgat ggagccagct	3120
cagacgggca cctggcacac atcttcagac atggggccaa ggagctgttc aacgaggatg	3180
tggaggaggt cacttaccga gccctgagaa acaaagactt ccaagaggtc acccttgaga	3240
agaacggaga ggtggtgtta cgctttgctg cagcctatgg ctttcgaaac atccagaaca	3300
tgatcctgaa gcttaagaag ggcaagttcc cattccactt tgtggaggtc ctgcctgtg	3360
ctggaggatg cttaaattggc agaggccaag ccagactcc agacggacat gcggataagg	3420
ccctgctgcg gcagatggaa ggcatttacg ctgacatccc tgtgcggcgt ccggagtcca	3480
gtgcacacgt gcaggagctg taccaggagt ggctggaggg gatcaactcc cccaaggccc	3540
gagagggtgct gcataccacg taccagagcc aggagcgtgg cacacacagc ctggacatca	3600
agtgggtgaag tcaggccagg gccttcacg tgcctctggg gccagagcca agagcctctc	3660
agtagaggga ggggctgccc tgagtggagt attaaagaca cttaagaaa	3709

<210> 46
 <211> 1840
 <212> DNA
 <213> Homo sapien

<400> 46	
cgcatctctc tctctgctaa gctttctcat gctagagaat gatgactaaa aaacagatga	60
tcccatcttc ttattcacca tcccctaaaa aaaacagact gccacgtgga atgtgtcagc	120
ctcacctggg ggagacaccc ttatctgacg acttaaaagg aaaaacttgg tgagccagcc	180
tccacacaaa cacacgcata gccttgcaaa aaacaaaaat ttattgggag aaagatggct	240
gatgggaggc aagatctgga ttcaggaatc ctcaatttta gctctcgctc tgtcaccaaa	300
accctccgct cctcccagag cctcagaggag tttccccttg tgtaaaccgg tatcgcgctc	360
cttcagctc tgacatcctg cggttctaga ataccattg attgaaagag ctgtggccaa	420
acaggaagga cagagtaaga gccgctgaca aaggattctg ccaagacaga atccaagggt	480
cacaggagag aggcttgggg cagtgaagc cctctcgggt cttggctttg gaagtgtgga	540
ctgtaccagc agatgggccc ttgggaagcc aacagggaac aattcaaggc tggagagaaa	600

82

```

aggccacctt tgaccagca gaaggcagag gcgggagagg agggtgagga tgtacaagga      660
gcacgcagg cgaggaaaca atggccagga cctaactgtg gtgggaactg cctttgtctc      720
cacacactcg caatcaacat gcgtatttgc tattctcaaa caactccctt ccacccctt      780
aggctgaaag gacaaactca tcagagttga ggagtaccta cctgatgaag ctgttcatgc      840
ttccagcatc aaactggacg tttactgtct gcccaaaagc cagtttccaa aaggtttgct      900
tccctctgtt cagtggattc ttgactacat ataggtcata tatttcaaaa aataatgcct      960
agctattttct actttgaaat catgactaaa gccaaaccac aaccacagca aagataacct     1020
aaggattttgt ttaccagaaa tacctacaaa aaagtttgca ggccgggccc gatggctcac     1080
acttgtaatc ccagcacttt gagaggccga gatgggcaga tcacctgagg tcaggggttc     1140
aagaccatcc tgccaacat gatgaaacct ggtctctact aaaaatacaa aaattaaccg     1200
gggtgtgggtg cgcatgcctg taatcccagc tactcaggag actgaggcgg aattgccaga     1260
accctggagg gcagaggttg cagtaagctg agatcgggcc actgcactcc agcctgggtg     1320
acagaatgag actcagtctc ccacctgcc aaaaagttag caatgtcggc agcaaagtgc     1380
aaacagagcc ggctctgggt gaaagcttta gttatgaaaa ataagaaaaa aaaaatccct     1440
aagcatttct taactagt tttaaagaaa acccctctc caaaggtagt tttctccttc     1500
aactttcttg actttgccag gctgtcagga tgcaagctac ctactctttt ctggatggat     1560
cgggacaaaag tttttaaaca aagactgtgt cctctccaca agcccttccc accaccacca     1620
ccccagcccc ccgcccgc atattacaggg gctgtgaaat ggctgttgag tgatggtcac     1680
acaaccagaa ggcagatgaa aaggtcaggc tctaagaggg gttcagcctc ccagtaccag     1740
cctgctgtgc ccaggggtgg gagtgctggg agaactgtct tccccagaca ggcagctatg     1800
ccgcctccca ctgcaaaggc accaaaggca acatccgtcc                        1840

```

<210> 47

<211> 1618

<212> DNA

<213> Homo sapien

<400> 47

```

cgtcgacggg atggagggtt ttggtgacag agcgagagct aaaattgagg attcctgaat      60
ccagatcttg cctcccatca gccatctttc tccagagcc tcgaggagt tccccttggtg     120
taaaccggta tcgcgtccct tccagctctg acatcctgcg gttctagaat acccattgat     180
tgaaagagct gtggccaaac aggaaggaca gagtaagagc cgctgacaaa ggattctgcc     240
aagacagaat cccaaggcca caggagagag gcttggggca gtgaaagccc tctcgggtct     300
tggctttgga agtgtggact gtaccagcag atgggccctt gggaagccaa cagggaaaca     360

```

83

ttcaaggctg	gagagaaaag	gccacctttg	accagcaga	aggcagaggc	gggagaggag	420
ggtgaggatg	tacaaggagc	atcgcaggcg	aggaaacaat	ggccaggacc	taactgtggt	480
gggaactgcc	tttgtctcca	cacactcgca	atcaacatgc	gtatttgcta	ttctcaaaca	540
actcccttcc	acccccttag	gctgaaagga	caaactcatc	agagttgagg	agtacctacc	600
tgatgaagct	gttcatgctt	ccagcatcaa	actggacggt	tactgtctgc	ccaaaagcca	660
gtttccaaaa	ggtttgcttc	cctctgttca	gtggattctt	gactacatat	aggatcatata	720
tttcaaaaaa	taatgcctag	ctatttctac	tttgaaatca	tgactaaagc	caaaccacaa	780
ccacagcaaa	gataacctaa	ggatttgttt	accagaaata	cctacaaaaa	agtttgcagg	840
ccgggcgcga	tggctcacac	ttgtaatccc	agcactttga	gaggccgaga	tgggcagatc	900
acctgaggtc	aggggttcaa	gaccatcctg	cccaacatga	tgaaacctgg	tctctactaa	960
aaatacaaaa	attaaccggg	tgtggtggcg	catgcctgta	atcccagcta	ctcaggagac	1020
tgaggcggaa	ttgccagaac	cctggagggc	agaggttgca	gtaagctgag	atcgggccac	1080
tgactccag	cctgggtgac	agaatgagac	tcagtctccc	accctgccaa	aaagtttgca	1140
atgtcggcag	caaagtgcaa	acagagccgg	ctctgggtga	aagctttagt	tatgaaaaat	1200
aagaaaaaaa	aaatccctaa	gcatttctta	acactagttt	taaagaaaac	cccctctcca	1260
aaggatatgt	tctccttcaa	ctttcttgac	tttgccaggc	tgtcaggatg	caagctacct	1320
actcttttct	ggatggatcg	ggacaaagtt	tttaaacaaa	gactgtgtcc	tctccacaag	1380
cccttcccac	caccaccacc	ccagccccc	gccgccataa	ttacaggggc	tgtgaaatgg	1440
ctgttgagtg	atggtcacac	aaccagaagg	cagatgaaaa	ggtcaggctc	taagaggggt	1500
tcagcctccc	agtaccagcc	tgctgtgccc	aggggtggga	gtgctgggag	aactgtcttc	1560
cccagacagg	cagctatgcc	gcctcccact	gcaaaggcac	caaaggcaac	atccgtcc	1618

<210> 48

<211> 3252

<212> DNA

<213> Homo sapien

<400> 48

agggggctct	gggcagacgt	aggcgccggg	ctgcaccttc	cctgcccgcc	tgtttgact	60
tctctgtgac	tgttggttc	aggtccttag	ccagcggagg	gctgactccg	gcctggctgc	120
cccctcccct	ttgcctcccg	tccctcccc	tggacactcc	aacccccatc	ccttcgtgtc	180
cctccccctc	ctctctctat	gctactccct	agccagaccc	cccgtctctc	cccggaacaa	240
tggaggggaa	tgtggctagg	ccagcctgac	gctgctgcgt	caacctctgg	ctcgttctcg	300
gcctctatcg	cttcaggga	cccttctatc	tgggctcagc	tgaccgcgcc	cgcctctcgg	360

cccacgtgcc	ggcagttctc	ccaggctggg	gctcagctgc	gctggagact	gggatgggtc	420
agctcagtg	ctctgtctgt	ttacaaagga	ggaacccagg	tgactaagc	aggcccccca	480
gtcctgggtg	caggcttggtg	acagctgctc	agggatccaa	acgtggcgct	tgagctctct	540
gcgatgtgtt	ccacagttcc	ctggaggcgg	acactgcgtg	agggtcagcc	atgtcatctg	600
agcctcccc	accaccacag	ccccccacc	atcaagcttc	agtcgggctg	ctggacaccc	660
ctcggagccg	tgagcgtca	ccatcccctc	tgcgcgga	cgtgggtcca	agcccactgc	720
ccactcgccg	gacgaggacc	ttctcggcga	cgggtgcggc	ttcacagggc	ccggtctaca	780
aaggagtctg	caaagtcttc	tgccgggtcca	agggccatgg	cttcattacc	ccagctgatg	840
gcgccccga	catcttctctg	cacatctctg	atgtggaagg	ggagtatgtc	ccagtggaag	900
gcgacgaggt	cacctataaa	atgtgctcca	tcccacccaa	gaatgagaag	ctgcaggccg	960
tggaggtcgt	catcactcac	ctggcaccag	gcaccaagca	tgagacctgg	tctggacatg	1020
tcatcagctc	ctaggagatg	gtggaagcac	cccttgtcct	gtgcttgtgg	gagactttgc	1080
ggggaggagg	cagcagacac	tggagatgac	attcttccac	acgagacggg	gcttcagccg	1140
ggcatggtcc	ctctcaagta	tctcctggag	gaaggggtat	ggggggcagg	tgtgggggtg	1200
ggggtgttcc	cggccatcag	cacagcctat	gaccattgca	acaacctctc	accatctgaa	1260
gagcattaaa	agcatttaaa	aaggagaggt	gccactggt	ggctgagtgg	aggttccaac	1320
cccatcccag	ggagtggatc	aagggtggta	tttctccagc	tgctcagaca	catggggtca	1380
accacagaa	tccctcttcc	tcttggagct	ggaggcccca	gattcccaga	tctggcccc	1440
tggcagcctg	acagggacct	tgctgactt	ctccaaggca	aatttccacc	taagtcccc	1500
ttgcgcctct	cctggggcct	gggcaaagca	gttttcta	tcttggcttg	gttggttcta	1560
ggggagctgg	cttgaagtgg	gtggggaaag	gcgggggtgg	cggctcttgg	attggacgga	1620
tgttgccctt	tgggtgcctt	gcagtgggag	gcggcatagc	tgctgtctg	gggaagacag	1680
ttctcccagc	actcccacc	ctgggcacag	caggctggta	ctgggaggct	gaacctctct	1740
tagagcctga	ccttttcatc	tgcttcttgg	ttgtgtgacc	atcactcaac	agccatttca	1800
cagccccctg	aattatggcg	gcggggggct	ggggtggtgg	tggtgggaag	ggcttgtgga	1860
gaggacacag	tctttgttta	aaaactttgt	cccgatccat	ccagaaaaga	gtaggtagct	1920
tgcatcctga	cagcctggca	aagtcaagaa	agttgaagga	gaaacatacc	tttggagagg	1980
gggttttctt	taaaactagt	gttaagaaat	gcttagggat	tttttttttc	ttatttttca	2040
taactaaagc	tttccccag	agccggctct	gtttgcactt	tgctgccgac	attgcaaact	2100
ttttggcagg	gtgggagact	gagtctcatt	ctgtcaccca	ggctggagtg	cagtggcccg	2160
atctcagctt	actgcaacct	ctgccctcca	gggttctggc	aattccgcct	cagtctcctg	2220

```

agtagctggg attacaggca tgcgccacca cacccggtta atttttgtat ttttagtaga 2280
gaccaggttt catcatgttg ggcaggatgg tcttgaaccc ctgacctcag gtgatctgcc 2340
catctcggcc tctcaaagtg ctgggattac aagtgtgagc catcgcgccc ggcctgcaaa 2400
cttttttgta ggtatttctg gtaaacaaat ccttaggtta tctttgctgt ggttgtgggt 2460
tggctttagt catgatttca aagtagaaat agctaggcat tttttttga aatatatgac 2520
ctatatgtag tcaagaatcc actgaacaga ggaagcaaa ccttttggaa actggccttt 2580
gggcagacag taaacgtcca gtttgatgct ggaagcatga acagcttcat caggtaggta 2640
ctcctcaact ctgatgagtt tgccttttca gcctaagggg gtggaaggga gttgtttgag 2700
aatagcaaat acgcatgttg attgcgagtg tgtggagaca aaggcagttc ccaccacagt 2760
taggtcctgg ccagtttggt tcctcgctg cgatgctcct tgtacgtcct caccctcctc 2820
tcccgctct gccttctgct gggccttctg tggccttttt ctctccagcc ttgaattggt 2880
ccctgttggc ttcccaaggg cccatctgct ggtacagtcc acacttccaa agccaagacc 2940
cgagagggtt ttactgccc caagcctctc tcctgtgacc ttgggattct gtcttggcag 3000
aatcctttgt cagcggtctc tgcctgtgcc ttctgtttg gccacagctc tttcaatcaa 3060
tggttattct agaaccgcag gatgtcagag ctggaaggga cgcgataccg gtttacacaa 3120
ggggaaactc ctcgaggctc tgggagggac ggagggtttt ggtgacagag cgagagctaa 3180
aattgaggat tcctgaatcc agatcttgcc tcccatcagc catctttctc ccaataaatt 3240
tttgttttgt gc 3252

```

```

<210> 49
<211> 3014
<212> DNA
<213> Homo sapien

```

```

<400> 49
ggcatgcgg gtgccctggc cccaggcca gcagctgggg cctttttgag gtgggggctc 60
cccggactca ggagttcctc cttggccgcg gctgtagctg agagaaggga caggaggtag 120
gatgccacct cctccagaag cgtggtgacc cgggctgtgt ggaaattcac tgggcctgat 180
ggagctggat gcagccttgg aaattctccc tttccctttg ggctcccgca gcaggcccc 240
cagtcctggg tgcaggcttg tgacagctgc tcagggatcc aaacgtggcg cttgagctct 300
ctgcgatgtg ttccacagtt ccctggaggc ggacactgcg tgagggtcag ccatgtcatc 360
tgagcctccc ccaccaccac agccccccac ccatcaagct tcagtcgggc tgctggacac 420
ccctcgagac cgtgagcgct caccatcccc tctgcgcggc aacgtgggtc caagcccact 480
gcccactcgc cggacgagga ccttctcggc gacggtgcgg gcttcacagg gccccgtcta 540

```

caaaggagtc	tgcaaatgct	tctgccggtc	caagggccat	ggcttcatta	ccccagctga	600
tggcgggcccc	gacatcttcc	tgcacatctc	tgatgtggaa	ggggagtatg	tcccagtggga	660
aggcgacgag	gtcacctata	aaatgtgctc	catcccaccc	aagaatgaga	agctgcaggc	720
cgtggagggtc	gtcatcactc	acctggcacc	aggcaccaag	catgagacct	ggctctggaca	780
tgtcatcagc	tcctaggaga	tgggtggaagc	accccttgctc	ctgtgcttgt	gggagacttt	840
gcgggggagga	ggcagcagac	actggagatg	acattcttcc	acacgagacg	gggcttcagc	900
cgggcatggg	ccctctcaag	tatctcctgg	aggaaggggt	atggggggca	gggtgtgggg	960
gtgggggtgtt	cccgccatc	agcacagcct	atgaccattg	caacaacctc	tcaccatctg	1020
aagagcatta	aaagcattta	aaaaggagag	gtgcccactg	gtggctgagt	ggagggtcca	1080
accccatccc	agggagtggga	tcaaggggtgg	tatttctcca	gctgctcaga	cacatgggct	1140
caaccacag	aatccctctt	cctcctggag	ctggaggccc	cagattccca	gatctggccc	1200
cctggcagcc	tgacagggac	cttgcgtagc	ttctccaagg	caaatttcca	cctaagtgcc	1260
ccttgcgctc	ctcctggggc	ctggggcaaag	cagttttcta	attcttggtc	tgggttggttc	1320
taggggagct	ggcttgaagt	gggtggggaa	aggcggggg	ggcggtcttt	ggattggacg	1380
gatgttgctc	tttgggtgct	ttgcagtggg	aggcggcata	gctgcctgtc	tggggaagac	1440
agttctccca	gcactcccac	ccctgggcac	agcaggctgg	tactgggagg	ctgaaccctc	1500
cttagagcct	gaccttttca	tctgccttct	ggttggtgta	ccatcactca	acagccattt	1560
cacagcccct	gtaattatgg	cggcgggggg	ctgggggtgg	ggtgggtggga	agggcttggtg	1620
gagaggacac	agtctttgtt	taaaaacttt	gtcccgatcc	atccagaaaa	gagtaggtag	1680
cttgcatcct	gacagcctgg	caaagtcaag	aaagtgaag	gagaaacata	cctttggaga	1740
gggggttttc	tttaaaacta	gtgttaagaa	atgcttaggg	attttttttt	tcttattttt	1800
cataactaaa	gctttcaccc	agagccggct	ctgtttgcac	tttgctgcgc	acattgcaaa	1860
ctttttggca	gggtgggaga	ctgagtctca	ttctgtcacc	caggctggag	tgagtgggcc	1920
cgatctcagc	ttactgcaac	ctctgccctc	cagggttctg	gcaattccgc	ctcagtctcc	1980
tgagtagctg	ggattacagg	catgcgccac	cacaccgggt	taatttttgt	atttttagta	2040
gagaccaggt	ttcatcatgt	tgggcaggat	ggtcttgaac	ccctgacctc	aggtgatctg	2100
cccatctcgg	cctctcaaag	tgctgggatt	acaagtgtga	gccatcgcg	ccggcctgca	2160
aacttttttg	taggtatttc	tggtaaacaa	atccttaggt	tatctttgct	gtggttggtg	2220
tttggtttta	gtcatgattt	caaagtagaa	atagctaggc	attatttttt	gaaatatatg	2280
acctatatgt	agtcaagaat	ccactgaaca	gagggaagca	aaccttttgg	aaactggctt	2340

87

ttgggcagac agtaaacgtc cagtttgatg ctggaagcat gaacagcttc atcaggtagg	2400
tactcctcaa ctctgatgag tttgtccttt cagcctaagg ggggtggaagg gagttgtttg	2460
agaatagcaa atacgcatgt tgattgcgag tgtgtggaga caaaggcagt tcccaccaca	2520
gttaggtcct ggccagtttg tttcctcgcc tgcgatgctc cttgtacgtc ctcaccctcc	2580
tctcccgccct ctgccttctg ctgggtcaaa ggtggccttt ttctctccag ccttgaattg	2640
ttccctgttg gcttcccaag ggcccatctg ctggtacagt ccacacttcc aaagccaaga	2700
cccagagagg ctttcactgc cccaagcctc tctcctgtga ccttgggatt ctgtcttggc	2760
agaatccttt gtcagcggct cttgctctgt ccttcctgtt tggccacagc tctttcaatc	2820
aatgggtatt ctagaaccgc aggatgtcag agctggaagg gacgcgatac cggtttacac	2880
aaggggaaac tcctcgaggc tctgggaggg acggagggtt ttggtgacag agcgagagct	2940
aaaattgagg attcctgaat ccagatcttg cctcccatca gccatcttcc tccaataaa	3000
tttttgtttt gtgc	3014

<210> 50
 <211> 2716
 <212> DNA
 <213> Homo sapien

<400> 50	
gctcaggacc ggtcgggaatt ccgggggcg g acagaacgga ttgcagggtc agccatgtca	60
tctgagcctc ccccaccacc acagccccc acccatcaag cttcagtcgg gctgctggac	120
accctcggga gccgtgagcg ctcaccatcc cctctgcgcg gcaacgtggg cccaagccca	180
ctgcccactc gccggacgag gaccttctcg gcgacggtgc gggcttcaca gggccccgtc	240
tacaaaggag tctgcaaatt cttctgccgg tccaagggcc atggcttcat taccacagct	300
gatggcggcc ccgacatctt cctgcacatc tctgatgtgg aaggggagta tgtccagtg	360
gaaggcgacg aggtcaccta taaaatgtgc tccatccac ccaagaatga gaagctgcag	420
gccgtggagg tcgtcatcac tcacctggca ccaggcacca agcatgagac ctggctctgga	480
catgtcatca gtccttagga gatgggtgga gcaccccttg tcctgtgctt gtgggagact	540
ttgcggggag gaggcagcag aactggaga tgacattctt ccacacgaga cggggcttca	600
gccgggcatg gtccctctca agtatctcct ggaggaaggg gtatgggggg cagggtgtggg	660
gtgtgggggtg ttcccgcca tcagcacagc ctatgaccat tgcaacaacc tctcaccatc	720
tgaagagcat taaaagcatt taaaaggag aggtgccac tgggtggctga gtggaggttc	780
caaccccatc ccagggagtg gatcaagggt ggtatttctc cagctgctca gacacatggg	840
ctcaaccac agaatccctc ttctcctgg agctggaggc ccagattcc cagatctggc	900

88

cccctggcag	cctgacaggg	accttgctg	acttctccaa	ggcaaatttc	cacctaagt	960
ccccttgccg	ctctcctggg	gcctgggcaa	agcagttttc	taattcttgg	cttgggttgg	1020
tctaggggag	ctggccttga	gtgggtgggg	aaaggcgggg	gtggcggtct	ttggattgga	1080
cggatgttgc	cttttggtgc	ctttgcagt	ggaggcggca	tagctgcctg	tctggggaag	1140
acagttctcc	cagcactccc	acccctgggc	acagcaggct	ggtactggga	ggctgaaccc	1200
ctcttagagc	ctgacctttt	catctgcctt	ctggttgtgt	gaccatcact	caacagccat	1260
ttcacagccc	ctgtaattat	ggcggcgggg	ggctgggggtg	gtgggtgggtg	gaagggtt	1320
tggagaggac	acagtctttg	tttaaaaact	ttgtcccgat	ccatccagaa	aagagtaggt	1380
agcttgcatc	ctgacagcct	ggcaaagtca	agaaagtga	aggagaaaca	tacctttgga	1440
gaggggggtt	tctttaaaac	tagtgtaag	aatgcttag	ggattttttt	tttcttattt	1500
ttcataacta	aagctttcac	ccagagccgg	ctctgtttgc	actttgctgc	cgacattgca	1560
aactttttgg	cagggtggga	gactgagtct	cattctgtca	cccaggctgg	agtgcagtgg	1620
cccgatctca	gcttactgca	acctctgccc	tccagggttc	tggcaattcc	gcctcagtct	1680
cctgagtagc	tgggattaca	ggcatgcgcc	accacaccgc	gttaattttt	gtatttttag	1740
tagagaccag	gtttcatcat	gttgggcagg	atggtcttga	acccctgacc	tcagggtgatc	1800
tgcccatctc	ggcctctcaa	agtgtctggga	ttacaagtgt	gagccatcgc	gcccggcctg	1860
caaacttttt	tgtaggattt	tctggtaaac	aatccttag	gttatctttg	ctgtgggtgt	1920
ggtttggtct	tagtcatgat	ttcaaagtag	aatagctag	gcattatttt	ttgaaatata	1980
tgacctatat	gtagtcaaga	atccactgaa	cagagggaag	caaacctttt	ggaaactggc	2040
ttttgggcag	acagtaaacg	tccagtttga	tgctggaagc	atgaacagct	tcattcaggta	2100
ggtactcctc	aactctgatg	agtttgtcct	ttcagcctaa	gggggtggaa	gggagttgtt	2160
tgagaatagc	aaatacgcat	gttgattgct	agtgtgtgga	gacaaaggca	gttcccacca	2220
cagttaggct	ctggccagtt	tgtttcctcg	cctgcgatgc	tccttgtaog	tcctcaccct	2280
cctctcccgc	ctctgccttc	tgttgggtca	aagggtggcct	ttttctctcc	agccttgaat	2340
tgttccctgt	tggcttccca	agggcccatc	tgctggtaca	gtccacactt	ccaaagccaa	2400
gacccgagag	ggctttcact	gccccaaagc	tctctcctgt	gaccttggga	ttctgtcttg	2460
gcagaatcct	ttgtcagcgg	ctottgctct	gtccttctctg	tttggccaca	gctctttcaa	2520
tcaatgggta	ttctagaacc	gcaggatgtc	agagctggaa	gggacgcgat	accggtttac	2580
acaaggggaa	actcctcgag	gctctgggag	ggacggaggg	ttttgggtgac	agagcgagag	2640
ctaaaattga	ggattcctga	atccagatct	tgcctcccat	cagccatctt	tctcccaata	2700
aatttttgtt	ttgtgc					2716

<210> 51
 <211> 3539
 <212> DNA
 <213> Homo sapien

<400> 51
 caatgagaag acttactttc ctggtaggca gcctgctttg ttttgcacag gaagatagat 60
 tttttttctc atctttctta taaacaacct catgcacatt ctgtgtttga gccaagacta 120
 gtcgaattcg aattcgaatt cgaagcggct ggcggacaga acggattgca gggtcagcca 180
 tgtcatctga gcctcccca ccaccacagc cccccacca tcaagcttca gtcgggctgc 240
 tggacacccc tcggagccgt gagcgctcac catccctct gcgcggcaac gtggtcccaa 300
 gccactgcc cactcgccgg acgaggacct tctcggcgtg agtctggagc aggtctggga 360
 cgcgttctca gatggatctc ttgcctgagg ccagagattc actgtgtctc tggccatttt 420
 tgggcctatg tcccggactg ccaaataggc agacctgggg ctgggccgtg ggctgagggg 480
 gagaaactag gaaggcgct tgtaaattca gggtagctgc cagcccagaa gaacctgtta 540
 ccagctggga gtggcaatgg gtgggcagtt gctggggctca gaatcgctgg gctttaacca 600
 cacatgagag tctggttccc tgtgagactg tgaattctgc agctgcaaaa tgggcacagc 660
 accagggctg caaaggacct acataatgcc tggcagatgc aggtctcagg ccttgcgagt 720
 ggaggatgct tgtgttatta atagtaaato attagtttct ggagggatgt tcagtgatgt 780
 ttttagttcaa ctccctcttt ccaggaggga aaacaggccc gagaaaattg gccagggtca 840
 cccaaagagt cttttccccc agcgtctttg tgagcggctc ccagggttga ggggtgtgca 900
 gtcctcagcc tgtgtgactg tacatggcgg tccctggccag ggcttgatgg ggcacagctg 960
 tgggcagggg ctggaatgag ctttgtgtgc ccaggagcgc tgatggggcc cctgacattt 1020
 cccctctct gtcaggacgg tgcgggcttc acaggggccc gtctacaaag gagtctgcaa 1080
 atgcttctgc cggccaagg gccatggctt cattaccca gctgatggcg gccccgacat 1140
 cttcctgcac atctctgatg tggaagggga gtatgtcca gtggaaggcg acgaggtcac 1200
 ctataaaatg tgctccatcc caccacaaga tgagaagctg caggccgtgg aggtcgtcat 1260
 cactcacctg gcaccaggca ccaagcatga gacctggtct ggacatgtca tcagctccta 1320
 ggagatggtg gaagcacccc ttgtcctgtg cttgtgggag actttgcggg gaggaggcag 1380
 cagacactgg agatgacatt cttccacacg agacggggct tcagccgggc atggctccctc 1440
 tcaagtatct cctggaggaa ggggtatggg gggcaggtgt ggggtgtggg gtgttcccg 1500
 ccatcagcac agcctatgac cattgcaaca acctctcacc atctgaagag cattaaggc 1560
 atttaaaaag gagaggtgcc cactggtggc tgagtggagg ttccaacccc atcccaggga 1620

gtggatcaag ggtggtat	ctccagctgc tcagacacat	gggctcaacc cacagaatcc	1680
ctcttcctcc tggagctgga	ggccccagat tcccagatct	ggccccctgg cagcctgaca	1740
gggaccttgc gtgacttctc	caaggcaaat ttccacctaa	gtgccccctg cgcctctcct	1800
ggggcctggg caaagcagtt	ttctaattct tggcttggtt	ggttctaggg gagctggctt	1860
gaagtgggtg gggaaaggcg	ggggtggcgg tctttggatt	ggacggatgt tgccttttgg	1920
tgcctttgca gtgggaggcg	gcatagctgc ctgtctgggg	aagacagttc tcccagcact	1980
cccacccctg ggcacagcag	gctggtactg ggaggctgaa	cccctcttag agcctgacct	2040
tttcatctgc cttctggttg	tgtgaccatc actcaacagc	catttcacag cccctgtaat	2100
tatggcggcg gggggctggg	gtggtggtgg tgggaagggc	ttgtggagag gacacagtct	2160
ttgtttaaaa actttgtccc	gatccatcca gaaaagagta	ggtagcttgc atcctgacag	2220
cctggcaaag tcaagaaagt	tgaaggagaa acataccttt	ggagaggggg ttttctttaa	2280
aactagtgtt aagaaatgct	tagggatttt ttttttctta	tttttcataa ctaaagcttt	2340
caccagagc cggctctgtt	tgcactttgc tgccgacatt	gcaaactttt tggcaggggtg	2400
ggagactgag tctcattctg	tcaccaggc tggagtgcag	tggcccgatc tcagcttact	2460
gcaacctctg ccctccaggg	ttctggcaat tccgcctcag	tctcctgagt agctgggatt	2520
acaggcatgc gccaccacac	ccggttaatt tttgtatttt	tagtagagac caggtttcat	2580
catgttgggc aggatggtct	tgaaccctg acctcagggtg	atctgcccat ctcggcctct	2640
caaagtgtg ggattacaag	tgtgagccat cgcgccggc	ctgcaaactt tttttagggt	2700
atttctggta aacaaatcct	taggttatct ttgctgtggt	tgtggtttgg ctttagtcat	2760
gatttcaaag tagaaatagc	taggcattat tttttgaaat	atatgacctt tatgtagtca	2820
agaatccact gaacagaggg	aagcaaacct tttggaaact	ggcttttggg cagacagtaa	2880
acgtccagtt tgatgctgga	agcatgaaca gcttcatcag	gtaggtactc ctcaactctg	2940
atgagtttgt cctttcagcc	taaggggggtg gaagggagtt	gtttgagaat agcaaatacg	3000
catgttgatt gcgagtgtgt	ggagacaaaag gcagttccca	ccacagttag gtcctggcca	3060
gtttgtttcc tcgcctgcga	tgctccttgt acgtcctcac	cctcctctcc cgcctctgcc	3120
ttctgctggg tcaaagggtg	cctttttctc tccagccttg	aattgttccc tgttggettcc	3180
ccaagggccc atctgctggt	acagtccaca cttccaaagc	caagaccoga gagggctttc	3240
actgccccaa gcctctctcc	tgtgaccttg ggattctgtc	ttggcagaat cctttgtcag	3300
cggtctctgc tctgtccttc	ctgtttggcc acagctcttt	caatcaatgg gtattctaga	3360
accgcaggat gtcagagctg	gaagggacgc gataccggtt	tacacaaggg gaaactcctc	3420

91

gaggctctgg gagggacgga gggttttggt gacagagcga gagctaaaat tgaggattcc 3480
 tgaatccaga tcttgccctcc catcagccat ctttctccca ataaattttt gttttgtgc 3539

<210> 52
 <211> 2810
 <212> DNA
 <213> Homo sapien

<400> 52
 caaggaattt tagggcggac gtcgggtagt ctttcccgcc tgcccacctg ttctctcggt 60
 gcaggggaag ctcttcccg cctccccgct gcttttgctt ttgggcggga gccctggacg 120
 ggggactcgc cgggcggtgg ggcgaggcgc ggggggcgtc gcacgcggga agcagagcca 180
 cccccgcgc ggcttgccg cgccagctgt ggctcgaggt tgggacgtca gccatgtcat 240
 ctgagcctcc cccaccacca cagcccccca cccatcaagc ttcagtcggg ctgctggaca 300
 cccctcggag ccgtgagcgc tcaccatccc cgtctacaaa ggagtctgca aatgcttctg 360
 ccggtccaag ggccatggct tcattacccc agctgatggc ggccccgaca tcttcctgca 420
 catctctgat gtggaagggg agtatgtccc agtggaaggc gacgaggtca cctataaaat 480
 gtgctccatc ccaccaaga atgagaagct gcaggccgtg gaggtcgtca tcactcacct 540
 ggcaccaggc accaagcatg agacctggtc tggacatgtc atcagctcct aggagatggt 600
 ggaagcacc cttgtcctgt gcttgtggga gactttgcgg ggaggaggca gcagacactg 660
 gagatgacat tcttccacac gagacggggc ttcagccggg catggtccct ctcaagtatc 720
 tcctggagga aggggtatgg ggggcagggtg tggggtgtgg ggtgttcccg gccatcagca 780
 cagcctatga ccattgcaac aacctctcac catctgaaga gcattaaaag catttaaaaa 840
 ggagaggtgc ccactggtgg ctgagtggag gttccaacct catcccaggg agtggatcaa 900
 ggggtggtatt tctccagctg ctccagacaca tgggctcaac ccacagaatc cctcttcctc 960
 ctggagctgg agggcccaga ttcccagatc tggccccctg gcagcctgac agggaccttg 1020
 cgtgacttct ccaaggcaaa tttccacctc agtgcccctt gcgcctctcc tggggcctgg 1080
 gcaaagcagt tttctaattc ttggcttggg tgggtctagg ggagctggct tgaagtgggt 1140
 ggggaaaggc ggggggtggc gtctttggat tggacggatg ttgccttttg gtgcctttgc 1200
 agtgggaggc ggcatagctg cctgtctggg gaagacagtt ctcccagcac tcccaccct 1260
 gggcacagca ggctggtact gggaggctga acccctctta gagcctgacc ttttcatctg 1320
 ccttctggtt gtgtgaccat cactcaacag ccatttcaca gcccctgtaa ttatggcggc 1380
 ggggggctgg ggtggtgggt gtgggaaggg cttgtggaga ggacacagtc tttgtttaaa 1440
 aactttgtcc cgatccatcc agaaaagagt aggtagcttg catcctgaca gcctggcaaa 1500

92

```

gtcaagaaag ttgaaggaga aacatacctt tggagagggg gttttcttta aaactagtgt 1560
taagaaatgc ttagggattt ttttttctt atttttcata actaaagctt tcaccagag 1620
ccggctctgt ttgcactttg ctgccgacat tgcaaacttt ttggcagggg gggagactga 1680
gtctcattct gtcaccaggt ctggagtgc gtggcccgat ctcagcttac tgcaacctct 1740
gccctccagg gttctggcaa ttccgcctca gtctcctgag tagctgggat tacaggcatg 1800
cgccaccaca cccggttaat ttttgtatth ttagtagaga ccaggtttca tcatgttggg 1860
caggatgggtc ttgaaccctt gacctcaggt gatctgccca tctcggcctc tcaaagtgtc 1920
gggattacaa gtgtgagcca tcgcgcccgg cctgcaaact tttttgtagg tatttctggg 1980
aaacaaatcc ttaggttatc tttgtgtggt ttgtggtttg gctttagtca tgatttcaaa 2040
gtagaaatag ctaggcatta ttttttgaaa tatatgacct atatgtagtc aagaatccac 2100
tgaacagagg gaagcaaacc ttttggaac tggttttgg gcagacagta aacgtccagt 2160
ttgatgctgg aagcatgaac agcttcatca ggtaggtact cctcaactct gatgagtgtg 2220
tcctttcagc ctaagggggg ggaagggagt tgtttgagaa tagcaaatac gcatgttgat 2280
tgcgagtgtg tggagacaaa ggcagttccc accacagtta ggtcctggcc agtttgtttc 2340
ctgcctcgcg atgtccttg tacgtcctca ccctcctctc ccgcctctgc cttctgctgg 2400
gtcaaagggtg gcctttttct ctccagcctt gaattgttcc ctggtggctt cccaagggcc 2460
catctgctgg tacagtccac acttccaaag ccaagaccog agagggttt cactgcccc 2520
agcctctctc ctgtgacctt gggattctgt cttggcagaa tcctttgtca gcggctcttg 2580
ctctgtcctt cctgtttggc cacagctctt tcaatcaatg ggtattctag aaccgcagga 2640
tgtcagagct ggaaggagcg cgataccggt ttacacaagg ggaaactcct cgaggctctg 2700
ggagggacgg agggtttttg tgacagagcg agagctaaaa ttgaggattc ctgaatccag 2760
atcttgccct ccatcagcca tctttctccc aataaatttt tgttttgtgc 2810

```

<210> 53

<211> 4725

<212> DNA

<213> Homo sapien

<400> 53

```

aatttgaaaa ccccatggga cctgagtcaa accgatccag ggctgctgct gtccctccc 60
gctgtgtggt gttgagtggc gtccctgagt aggtctctcc ctacttcac tcatacccc 120
atcaccagcc ctgcgcaggt gccaaagtgt gccaccag ctgccccact ttgcagattc 180
ggaaactgag ccagagatg ggaaaggact taccaaggg cctctgactg ctggactgtt 240
tctctttccc ctgaacctcg agctgaggcc acgttagagc cgacaccttc atattcaatg 300

```

ttcagtgatt attgtgtaca gcgagtgtcc tgctcttgaa aaattgtcaa aagcacagca	360
gcagctaggg aattcccgcc ctccagctgg atttaccggg acatactagg ggtcacagaa	420
ggaatttgat ttcataacag cgtgggtgacc cgggctgtgt ggaaattcac tgggcctgat	480
ggagctggat gcagccttgg aaattctccc tttccctttg ggctcccgca gcaggccccc	540
cagtccctggg tgcaggcttg tgacagctgc tcagggatcc aaacgtggcg cttgagctct	600
ctgcgatgtg ttccacagtt ccctggaggc ggacactgcg tgagggtcag tttttgcccc	660
tgcttccctg tgggtgtctgg ctcccagcag cgtcaggag ggtccgtgga gtggctgagt	720
ttggatctag gtggttggca ttaaaatctc cctggctttg ggttttcttt tttgagacag	780
agtcttgctc tgtagcccag gctggagtgc agtgggtgca tctcagctca ctggaacctc	840
cgcctcccag gttcaagcaa ttctcatgcc tcagccttca agtagatggg attacagggtg	900
cctgccacca tgcccagcta atttttgtat tcgtattaga gactgggttt ccccatgttg	960
gccaggctag tctcgaactc ctgacattaa gtgatccgcc tgccctcgcc tcccagagt	1020
ctgggattgc aggtgtgagc cactgcgccc ggcccttggc tttgggtttt ctgctcacct	1080
tcctccttcc ctcccacaaa tactatgtgt cccaaatgtg ccgggatcca tgccttgtgc	1140
tggggacaca gaggggaccg atgccatcct ggtcatgggg gaggatgtgg cattttcatg	1200
aagagctatg atcgaggggg cccacagggt gcttcaggag ccagagtag gcatctgacc	1260
ctgccaggcg tgctcccaga caaggtgaca atgggtgctga ggtgcccgag aggaatgagc	1320
ctcagatact ggactgtcag acgtggacat gggggtaagc acaggggtccc taccttcaa	1380
ggcacgagcc cctgacagcc tcttctcttt ccagggtcagc catgtcatct gagcctcccc	1440
caccaccaca gccccccacc catcaagctt cagtggggct gctggacacc cctcggagcc	1500
gtgagcgctc accatcccct ctgcgcggca acgtgggtccc aagccactg cccactcgcc	1560
ggacgaggac cttctcggcg tgagtctgga gcaggctggg gacgcgttct cagatggatc	1620
tcttgccctga ggccagagat tcactgtgtc tctggccatt tttgggccta tgtcccggac	1680
tgccaaatag gcagacctgg ggctgggccc tgggctgagg gtgagaaact aggaaggcgc	1740
cttgtaaatt cagggtactg cccagcccag aagaacctgt taccagctgg gagtggcaat	1800
gggtgggcag ttgctggggt cagaatcgct gggctttaac cacacatgag agtctggttc	1860
cctgtgagac tgtgaattct gcagctgcaa aatgggcaca gcaccagggc tgcaaaggac	1920
ccacataatg cctggcagat gcaggctctca ggccttgcca gtggaggatg cttgtgttat	1980
taatagtaaa tcattagttt ctggagggat gttcagtgat gttttagttc aactccctct	2040
ttccaggagg gaaaacaggc ccgagaaaat tggccagggt cacccaaaga gtcctttccc	2100
ccagcgtctt tgtgagcggc tcccagggtt gagggttgtg cagtcctcag cctgtgtgac	2160

tgtacatggc ggtcctggcc agggcttgat ggggcacagc tgtgggcagg ggctggaatg 2220
agcttttgtgt gccagggagc gctgatgggg cccctgacat ttccccctct ctgtcaggac 2280
gggtgcgggct tcacagggcc ccgtctacaa aggagtctgc aaatgcttct gccggtccaa 2340
gggccatggc ttcattaccc cagctgatgg cggccccgac atcttcctgc acatctctga 2400
gtgagtcagc ggcaggctctg ggtggccttg ctgggggtgg cgcttccttg gccctggcac 2460
caggcaccaa gcatgagacc tggctctggac atgtcatcag ctccctaggag atggtggaag 2520
cacccttctgt cctgtgcttg tgggagactt tgcggggagg aggcagcaga cactggagat 2580
gacattcttc cacacgagac ggggcttcag ccgggcatgg tccctctcaa gtatctcctg 2640
gaggaagggg tatggggggc aggtgtgggg tgtgggggtgt tcccggccat cagcacagcc 2700
tatgaccatt gcaacaacct ctaccatct gaagagcatt aaaagcattt aaaaaggaga 2760
ggtgcccact ggtggctgag tggaggttcc aaccccatcc cagggagtgg atcaagggtg 2820
gtattttctcc agctgctcag acacatgggc tcaaccacaa gaatccctct tcctcctgga 2880
gctggaggcc ccagattccc agatctggcc ccctggcagc ctgacagggg ccttgctga 2940
cttctccaag gcaaatttcc acctaatgct cccttgccgc tctcctgggg cctgggcaaa 3000
gcagttttct aattcttggc ttgggttggt ctaggggagc tggcttgaag tgggtgggga 3060
aaggcggggg tggcggtctt tggattggac ggatgttgcc ttttggtgcc tttgcagtgg 3120
gaggcgggcat agctgcctgt ctggggaaga cagtctctcc agcactccca cccctgggca 3180
cagcaggctg gtactgggag gctgaacccc tcttagagcc tgaccttttc atctgccttc 3240
tgggttgtgt accatcactc aacagccatt tcacagcccc tgtaattatg gcggcggggg 3300
gctgggggtg tgggtggtgg aagggtctgt ggagaggaca cagtctttgt ttaaaaactt 3360
tgtcccgatc catccagaaa agagtaggta gcttgcatcc tgacagcctg gcaaagtcaa 3420
gaaagttgaa ggagaaacat acctttggag agggggtttt ctttaaaaact agtgtaaga 3480
aatgcttagg gatTTTTTTT ttcttatttt tcataactaa agctttcacc cagagccggc 3540
tctgtttgca ctttgctgcc gacattgcaa actttttggc aggggtgggag actgagtctc 3600
attctgtcac ccaggctgga gtgcagtggc ccgatctcag cttactgcaa cctctgccct 3660
ccagggttct ggcaattccg cctcagtctc ctgagtagct gggattacag gcatgcgcca 3720
ccacaccgg ttaatttttg tatTTTTtagt agagaccagg ttcatcatg ttgggcagga 3780
tggctctgaa cccctgacct cagggtgatct gccatctcg gcctctcaaa gtgctgggat 3840
tacaagtgtg agccatcgcg cccggcctgc aaactTTTTT gtaggtattt ctggtaaaca 3900
aatccttagg ttatctttgc tgtggttgtg gtttggcttt agtcatgatt tcaaagtaga 3960

95

```

aatagctagg cattatTTTT tgaatatat gacctatatg tagtcaagaa tccactgaac 4020
agaggggaagc aaacctTTTT gaaactggct tttgggcaga cagtaaacgt ccagtttgat 4080
gctggaagca tgaacagctt catcaggtag gtactcctca actctgatga gtttgcctt 4140
tcagcctaag ggggtggaag ggagttgttt gagaatagca aatacgcag ttgattgcga 4200
gtgtgtggag acaaaggcag ttcccaccac agttaggtcc tggccagttt gtttcctcgc 4260
ctgcgatgct ccttgtacgt cctcaccctc ctctcccgcc tctgccttct gctgggtcaa 4320
aggtggcctt tttctctcca gccttgaatt gttccctgtt ggcttcccaa gggcccatct 4380
gctggtacag tccacacttc caaagccaag acccgagagg gctttcactg cccaagcct 4440
ctctcctgtg accttgggat tctgtcttgg cagaatcctt tgtcagcggc tcttgctctg 4500
tccttcctgt ttggccacag ctctttcaat caatgggtat tctagaaccg caggatgtca 4560
gagctggaag ggacgcgata ccggtttaca caaggggaaa ctctcgagg ctctgggagg 4620
gacggaggggt tttggtgaca gagcgagagc taaaattgag gattcctgaa tccagatctt 4680
gcctcccatc agccatcttt ctccaataa atttttgttt tgtgc 4725

```

<210> 54

<211> 1147

<212> DNA

<213> Homo sapien

<400> 54

```

gaggatcccc atcgaggcsa tgtgaagaca gccagccgac attcttccaa kkggaagatg 60
tgcatagaaa cagatacaaa caaatcacat atatggctca aggtcttgct gtgagagaga 120
tgactgggat gactaagttt aaacctattt tcatttctgc aacttctgag attttgtctc 180
aaaaatgcat taatacaaat gtgctTTTT tgtctctctc agataatcat ggacatattg 240
atccctctct gagactcatc tgggatttgg ctttccttgg aagctcttac gtgatgtggg 300
aatgacaac acaggtgtca cactactact tggcacagct cactagtgtg aggagtgga 360
aaactaatga tgatacaatt gatTTtgatt atactgttct acttcatgaa ttatcaacac 420
aggaaataat tccctgtcgc attcacttgg tctggtaccc tggcaaacct cttaaagtga 480
agtaccactg tcaagagcta cagacaccag aagaagcctc cggaactgaa gaaggatcag 540
ctgtagtacc aacagagctt agtaatttct aaaaagaaaa aatgatcttt ttccgacttc 600
taaacaagtg actatactag cataaatcat tcttctagta aaacagctaa ggtatagaca 660
ttctaataat ttgggaaaac ctatgattac aagtaaaaaac tcagaaatgc aaagatgttg 720
gttttttgtt tctcagctctg ctttagcttt taactctgga agcgcatgca cactgaactc 780
tgctcagtg taaacagtca ccagcaggtt cctcaggggt tcagccctaa aatgtaaaac 840

```

96

ctggataatc agtgtatggt gcaccagaat cagcattttt tttttaactg caaaaaatga	900
tgggtctcatc tctgaattta tttttctcat tcttttgaac atactatagc taatatattt	960
tatgttgcta aattgcttct atctagcatg ttaaacaaag ataataact ttcgatgaaa	1020
gtaaattata ggaaaaaaat taactgtttt aaaaagaact tgattatggt ttatgatttc	1080
aggcaagtat tcattttttaa cttgctacct acttttaaact aaatgtttac atttctaaat	1140
aatgaaa	1147

<210> 55
 <211> 1415
 <212> DNA
 <213> Homo sapien

<400> 55	
caactgcctg cctactgaaa aatatccaat gcactgggtt tagaaacagc gtgtatagct	60
tcatgtggta tatatgaaga atgtgggtgt tggagtcagt agatctgggt tggaattccc	120
atttgtccct tcctagctgt aatctcaaat tacctctcct gtaaaaccag gtcataata	180
cacagattgt aaaggagaat tagatgagaa catgtagaag gtgcctagca cagaaccag	240
cccatgggtg accttctaga attactgatt gccctcccc ttgtcttatt gcctgtcata	300
tcagggtctaa tatatcagct tgaggatat atacagtatc tagatcctta aggttggtta	360
aagaatctcc ctattcccac agacttcggt ctccatcaga aaaagtgaca gacattcaac	420
cagactttca tataactcag tgacaagaca aaaatgagga tccatgactt ccatatcttg	480
aagcaaaata ccaccaaaaa ccgtgaggct gaaatcgaaa aggagtagg agacacaagg	540
caccattca aaacaagaat gtattgtata gttattacac taaataccac catatttatt	600
acactaactc ttttttctcc ctctagaaaa actaatgatg atacaattga ttttgattat	660
actgttctac ttcatgaatt atcaacacag gaaataattc cctgtcgcat tcacttggtc	720
tggtagcctg gcaaacctct taaagtgaag taccactgtc aagagctaca gacaccagaa	780
gaagcctccg gaactgaaga aggatcagct gtagtaccac cagagcttag taatttctaa	840
aaagaaaaaa tgatcttttt ccgacttcta aacaagtgac tatactagca taaatcattc	900
ttctagtaaa acagctaagg tatagacatt ctaataattt gggaaaacct atgattacaa	960
gtaaaaactc agaaatgcaa agatgttgggt tttttgtttc tcagtctgct ttagctttta	1020
actctggaag cgcattgaca ctgaactctg ctgagtgtta aacagtcacc agcaggttcc	1080
tcagggtttc agccctaaaa tgtaaaacct ggataatcag tgtatgttgc accagaatca	1140
gcattttttt ttttaactgca aaaaatgatg gtctcatctc tgaatttata tttctcattc	1200
ttttgaacat actatagcta atatatttta tgttgctaaa ttgcttctat ctgcatggtt	1260

97

aaacaaagat aatatacttt cgatgaaagt aaattatagg aaaaaaatta actgttttaa 1320
 aaagaacttg attatgtttt atgatttcag gcaagtattc atttttaact tgctacctac 1380
 ttttaaataa atgtttacat ttctaaataa tgaaa 1415

<210> 56
 <211> 2201
 <212> DNA
 <213> Homo sapien

<400> 56
 caggatcaga tacgcacttc tgcttcttct acaggagaga agagaatagt tttgatagag 60
 accaggggaag aagccacctg tgccaaagca tggatgggga ccagacagaa gagagaatga 120
 tgaagatgat ggtgcatcag agacctttgc ccagccagc actacttct atgagctcca 180
 acactttccc ttcaagatct actaagccaa gtcccatgaa ccctctccca tcctctcaca 240
 tgcttgagc attctcagaa agtaacagca gttttccaca gagtgcctcc ctgccaccat 300
 acttctctca aggccctagc aacagaccac ctatcagagc cgaaggcaga aacttcccct 360
 tgccacttcc aaacaaacct cggcccccat ccccgcgga ggaagagaat tcattaaatg 420
 aagagtggta cgtttcttat attaccgcag cagaggcaga agctgctctt agaaagataa 480
 accaggatgg cacatttctg gtcagagaca gctctaaaaa aacaacaacc aatccatatg 540
 tcctcatggg gttgtacaaa gataaagttt acaacatcca gatccgttat cagaaggaaa 600
 gtcaagttta cttgttggga actggactcc gagggaaaga ggactttctg tctgtgtcag 660
 atattattga ctacttcagg aaaatgccac ttctgctcat tgatgggaaa aaccgaggtt 720
 ccagatacca gtgcacatta acgcatgctg cagggtaccc atagcaagtt atagccgagc 780
 aatgaaccg tcctcctgcc tctgttgcca acacgagatc aatcagcctt ggtcaatgga 840
 caaacactta ggactgaact gaacccctcc ccatgaacac aagggtttta tcctttcctt 900
 taaaaacagt gtttgaaatg aagactgtca actatcccat aatttattta ttcttcttca 960
 atgtttgtaa agtgcagag tcatgttcac acttgaagtc tagtagtgca ctgtaataat 1020
 tcatttttta aaagattatt taatgcccc ttcaaaatac agtagtttac acagctacag 1080
 aaacaatttg gggcaagttt taaaacactg aaacagtaat agttattggg gtcacataaa 1140
 actgatttgt tttttacagc caaacctctg tcagtcagag gcattcatta gttttataca 1200
 tgtaatttga aaatcactaa acctcgtttt ctcagcagca ataatttaag aggcttcaaa 1260
 aatataatth cactcttatt tagtattttt tcctgggggc atttttacgt aattttttta 1320
 tgaaaagaca aatgcatggt gagataactt ctgggattaa aatagtcttt tgctttactt 1380
 ttttggttct ctaaaacaac ttatttgact tttagtccat actgttatat ttttgtctta 1440

98

```

aagaaaattt aaactacaaa taccaaaaga aaacatttta aatttaggga tgagactttg 1500
gtgtatcgtg ggtctagggt taatgaacac atctgggggt aagttggcat ttcttcacat 1560
ctccacaccc acaccaacca tcacagcccc ccaccaacct tctcccaacc ccaaaagcat 1620
tgtccagga tatagatttt accaaaggct tcctgggaag acgagggagc aacactttag 1680
attaaatgtg atcagacttt cctattagat atggctcttc tgtctcttgt tatccccctg 1740
acagctctgc cataaagtcc cttctcctca tccttcccaa acaggctgta taagtgtttt 1800
gaggtaatta aactctttcc tccagtttac aaatatcact taacaaaaaa tataggcatt 1860
cagccagatt aaaaaactgg tattcagcca aatagtgaca atcagttgtt cttcaagttt 1920
ttccctttgg gaccttggtt gttattgcac aacttttatt agcaacaatt tttggcgtct 1980
ctgcttaatc tacaagtttt cgaaatggaa aagagtatct tgcagcttca ttttcatgag 2040
ctaataaaag ggggtattgga aggaatctaa gaagtcacca ttttaaaact gatgatatgt 2100
taaaataagg agtatgcaga aggtagagac ttttaactga tgataaaaaat ggtgtttcac 2160
aaaatctcat ccttaacaac cagaagttct cagtttaggg t 2201

```

<210> 57

<211> 2435

<212> DNA

<213> Homo sapien

<400> 57

```

ccccttcag tccctgctcc ttcaatagac agaagcacga aacctcccct agatcgttca 60
ttagctccgt ttgatagaga acccttcaca ctaggggagg ttgcgtgctt ctgtctattg 120
aaggagcacg aaacctcccc tagatcgttc attagctccg tttgatagag aaccttcac 180
actaggaaag aaaccacat tttctgacaa gccctcgatt ccagcgggaa ggtcactcgg 240
ggagcattta cccaagattc aaaagcctcc tttaccaccg accacggaaa gacatgaaag 300
gagcagcccc ctgccagga agaagccacc tgtgccaaag catggatggg gaccagacag 360
aagagagaat gatgaagatg atgtgcatca gagacctttg cccagccag cactacttcc 420
tatgagctcc aacactttcc cttcaagatc tactaagcca agtcccatga accctctccc 480
atcctctcac atgcctggag cattctcaga aagtaacagc agttttccac agagtgcctc 540
cctgccacca tacttctctc aaggccctag caacagacca cctatcagag ccgaaggcag 600
aaacttcccc ttgccacttc caaacaacc tcggcccca tccccgcgg aggaagagaa 660
ttcattaaat gaagagtggg acgtttctta tattaccga ccagaggcag aagctgctct 720
tagaaagata aaccaggatg gaaacatttc tcacatttat gtcttctaac acacgtgggt 780
aaagaatcag taactgccaa tgaatgagat agcataaaat tttgactttc ataaaacaaa 840

```

99

```

acttgtaa at catttctgaa atttataagg tttccagctg acaatcttct gccaaagttaa 900
catttcaaaa atgaaaaaca aaatcaatgt gtatttttaga gtatcatcta gatctgcgcc 960
cctttgtaaa tggattctat ggtgggttaag tcatctagaa acaccctggg agcaagggtc 1020
catcctacta ctctctaatt acatctgac tgctttattt ttgtgtcctc ttataactag 1080
ctgtttctaa ctggggggcg gttatggaat gtttcttagt gacagctggg gcctgccaat 1140
ctcattcatt gttttctact tctgcacgtg catcagagac ctttgcccca gccagcacta 1200
cttctatga gctccaacac tttcccttca agatctacta agccaagtcc catgaaccct 1260
ctccatcct ctcacatgcc tggagcattc tcagaaagta acagcagttt tccacagagt 1320
gcctccctgc caccatactt ctctcaaggc cctagcaaca gaccacctat cagagccgaa 1380
ggcagaaaact tccccttgcc acttccaaac aaacctcggc ccccatcccc cgcgaggagaa 1440
gagaattcat taaatgaaga gtggtacgtt tcttatatta cccgaccaga ggcagaagct 1500
gctcttagaa agataaacca ggatggcaca tttctgggtca gagacagctc taaaaaaca 1560
acaaccaatc catatgtcct catggtgttg taaaaagata aagtttaca catccagatc 1620
cgttatcaga aggaaagtca agtttacttg ttgggaactg gactccgagg gaaagaggac 1680
tttctgtctg tgtcagatat tattgactac ttcaggaaaa tgccacttct gtcattgat 1740
gggaaaaacc gaggttccag ataccagtgc acattaacgc atgctgcagg gtacccatag 1800
caagttatag ccgagcaaat gaaccgtcct cctgcctctg ttgccaacac gagatcaatc 1860
agccttggtc aatggacaaa cacttaggac tgaactgaac ccctcccat gaacacaagg 1920
gttttatcct ttcctttaaa aacagtgtt gaaatgaaga ctgtcaacta tcccataatt 1980
tatttattct tcttcaatgt ttgtaaagt catgagtcatt gttcacactt gaagtctagt 2040
agtgcactgt aataattcat tttttaaaag attatttaatt gccatttca aaatacagta 2100
gtttacacag ctacagaaac aatttggggc aagttttaaa aactgaaac agtaatagtt 2160
attggtgtca cataaaactg atttgttttt tacagccaaa cctctgtcag tcagaggcat 2220
tcattagttt tatacatgta atttgaaaat cactaaacct cgttttctca gcagcaataa 2280
tttaagaggc ttcaaaaata taatttcaact cttatttagt attttttctt gggggcattt 2340
ttacgtaatt tttttatgaa aagacaaatg catgttgaga taacttctgg gattaaaata 2400
gtcttttgct ttaaaaaacc tcgccgcgac cagcgc 2435

```

<210> 58

<211> 3612

<212> DNA

<213> Homo sapien

<400> 58

100

cgaactccgc cccctcttct cgcgtaggc ccagctccct gaagcggctg tttcgagcca	60
cgcgcccatc gggtagcgag gcacgcgcgc ggcgtcacgt gcgtttcgcg gcgagcggaa	120
atgacgcgag ttgtgtgagc cgccagtatg gccgggctat ggcggcgagc actggctacg	180
tgcgactgtg gggagcggcg cgggtgctggg tgctgcggcg gccgatgctg gccgccgcgc	240
gggggcgggt tcccactgca gcaggagcgt ggttgctccg aggccagcgg acctgcgacg	300
cctctcctcc ttgggcactg tggggccgag gcccggaat tgggggcca tggcgggggt	360
tttggaagc gagcagccgc ggcggaggcg cattctcggg gggcgaggac gcctccgagg	420
gcggcgcgga ggaaggagcc ggcggcgcg ggggcagcgc gggcgccggg gaaggcccgg	480
tcataacggc gctcacgccc atgacgatcc ccgatgtgtt tccgcacctg ccgctcatcg	540
ccatcacccg caaccgggtg tccccgcgt ttatcaagat tatcgagggt aaaaataaga	600
agttgggtga gctgctgaga aggaaagttc gtctcgccca gccttatgtc ggcgtctttc	660
taaagagaga tgacagcaat gagtcggatg tggtcgagag cctggatgaa atctaccaca	720
cggggacgtt tgcccagatc catgagatgc aggaccttgg ggacaagctg cgcgatgatc	780
tcattgggaca cagaagagtc catatcagca gacagctgga ggtggagccc gaggagccgg	840
aggcgagaaa caagcacaag cccgcagga agtcaaagcg gggcaagaag gaggcgagg	900
acgagctgag cgccaggcac ccggcggagc tggcgatgga gccaccacct gagctcccg	960
ctgaggtgct catggtggag gtagagaacg ttgtccacga ggacttccag gtcacggagg	1020
agggtgaaagc cctgactgca gagatcgtga agaccatccg ggacatcatt gccttgaacc	1080
ctctctacag ggagtcagtg ctgcagatga tgcaggctgg ccagcgggtg gtggacaacc	1140
ccatctacct gagcgacatg ggcgcgcgc tcaccggggc cgagtcccat gagctgcagg	1200
acgtcctgga agagaccaat attcctaagc ggctgtacaa ggccctctcc ctgctgaaga	1260
aggaatttga actgagcaag ctgcagcagc gcctggggcg ggaggtggag gagaagatca	1320
agcagacca ccgtaagtac ctgctgcagg agcagctaaa gatcatcaag aaggagctgg	1380
gcctggagaa ggacgacaag gatgccatcg aggagaagtt ccgggagcgc ctgaaggagc	1440
tcgtggtccc caagcacgtc atggatgttg tggacgagga gctgagcaag ctgggcctgc	1500
tggacaacca ctctcggag ttcaatgtca cccgcaacta cctagactgg ctcacgtcca	1560
tcccttgggg caagtacagc aacgagaacc tggacctggc gcgggcacag gcagtgtgg	1620
aggaagacca ctacggcatg gaggacgtca agaaacgcat cctggagttc attgccgtta	1680
gccagctccg cggctccacc cagggaaga tcctctgctt ctatggcccc cctggcgtgg	1740
gtaagaccag cattgtctgc tccatcgccc gcgccctgaa ccgagagtag ttccgcttca	1800
gcgtcggggg catgactgac gtggctgaga tcaagggcca caggcggacc tacgtgggcg	1860

ccatgcccgg gaagatcatc cagtgtttga agaagaccaa gacggagaac cccctgatcc	1920
tcatcgacga ggtgcgtgcg gagaggcccg gagaccccct aaccgcgcaa tgccgtgtgc	1980
aagatgaagc cgagccttcc gggtcctagg atccccgagc caactcaggg gccatcgccg	2040
tgatgctggc cctgatggca cggacgggtc cctcccaccc acagccacgc ccagcaggct	2100
gcccgtccaa ggcagggtgt gggtgccgct ctactcggg gtctcccag ccctgaccct	2160
gggggtctccc cactgggagg gctggccagg ggcacggcga ggggcaccac ctgctgtctg	2220
ctgccccagg tggacaagat cggccgaggc taccaggggg acccgctctc ggactgctg	2280
gagctgctgg acccagagca gaatgccaac ttcttgacc actacctgga cgtgccctg	2340
gacttgtcca aggtgggggg cctgtctggg gcctgggctt gctggggagg gtaagcgcc	2400
agtccccag caccttgccg ccgccccag gtgctgttca tctgcacggc caacgtcacg	2460
gacaccatcc ccgagccgct gcgagaccgt atggagatga tcaacgtgtc gggctacgtg	2520
gccaggaga agctggccat tgcggagcgc tacctggtgc ccaggctcg cgccctgtgt	2580
ggcttgatg agagcaaggc caagctgtca tcggacgtgc tgacgctgct catcaagcag	2640
tactgccgag agagcgggtgt ccgcaacctg cagaagcaag tggagaaggt gttacggaaa	2700
tcggcctaca agattgtcag cggcgaggcc gagtccgtgg aggtgacgcc cgagaacctg	2760
caggacttcg tggggaagcc cgtgttcacc gtggagcgca tgtatgacgt gacaccgcc	2820
ggcgtggtca tggggctggc ctggaccgca atgggaggct ccacgctgtt tgtggagaca	2880
tccctgagac ggccacagga caaggatgcc aagggtgaca aggatggcag cctggagggtg	2940
acaggccagc tgggggaggt gatgaaggag agcgcccgca tagcctacac cttcgccaga	3000
gccttcctca tgcagcacgc ccccgccaat gactacctgg tgacctaca catccacctg	3060
catgtgcccg agggcgccac cccaaggac ggccaagcg caggctgcac catcgtcacg	3120
gccctgctgt ccctggccat gggcaggcct gtccggcaga atctggccat gactggcgaa	3180
gtctccctca cgggcaagat cctgcctgtt ggtggcatca aggagaagac cattgcggcc	3240
aagcgcgagc ggggtgacgtg catcgtcctg ccagccgaga acaagaagga cttctacgac	3300
ctggcagcct tcatcaccga gggcctggag gtgcacttcg tggaacta ccgggagatc	3360
ttcgacatcg ccttcccga cgagcaggca gaggcgctgg ccgtggaacg gtgacggcca	3420
ccccgggact gcaggcggcg gatgtcaggc cctgtctggg ccagaactga gcgctgtggg	3480
gagcgcgccc ggacctggca gtggagccac cgagcgagca gctcggtcca gtgaccaga	3540
tcccagggac ctacgtcggc ttaatcagag tgtggcatag aagctattta atgattaaag	3600
tcatttgcag ta	3612

<210> 59
<211> 2744
<212> DNA
<213> Homo sapien

<400> 59
tcctgggttc acgccattct cctgcctcag cctcccagagt agctgggact acaggcgccc 60
gccatcacgc ccgcctaatt tttgtathtt ttggtagaga cggggtttca ctgtgttagc 120
caggaggggtc tcgatctcct gacctcatga tccacccgcc tcggcctccc aaagtgcttg 180
gattacaggc gtgagccacc acgcccggcc aagacttggc ctcttaaaaa atggaggaaa 240
gaaaacacag tctctttttt tcccctgaga tagcttatcc taaggacaag gctctgaggg 300
gtaaggagac agggtttggc tttagttttg actctcttgg tcctgtgtcc ctgggactgg 360
tgagtccacc agggactgca ctctccatc tcccaggccg tccgtggcca gggccagttg 420
gcgaggggga gggagtgtgat ggagctggc ggtggggctc ccagggccat gctcccaccc 480
agcccccccc ctgctggagc actgctgggc cccagcgacc ctgccccata cttggcgccc 540
tgactgcaga gatcgtgaag accatccggg acatcattgc cttgaaccct ctctacaggg 600
agtcagtgtc gcagatgatg caggctggcc agcgggtggc ggacaacccc atctacctga 660
gcgacatggg cgccgcgctc accggggccg agtcccatga gctgcaggac gtccctggaag 720
agaccaatat tcctaagcgg ctgtacaagg ccctctccct gctgaagaag gaatttgaac 780
tgagcaagct gcagcagcgc ctggggcggg aggtggagga gaagatcaag cagacccacc 840
gtaagtacct gctgcaggag cagctaaaga tcatcaagaa ggagctgggc ctggagaagg 900
acgacaagga tgccatcgag gagaagtcc gggagcgcct gaaggagctc gtggtcccca 960
agcacgtcat ggatgttgtg gacgaggagc tgagcaagct gggcctgtct gacaaccact 1020
cctcggagtt caatgtcacc cgcaactacc tagactggct cacgtccatc ccttggggca 1080
agtacagcaa cgagaacctg gacctggcg gggcacaggc agtgctggag gaagaccact 1140
acggcatgga ggacgtcaag aaacgcatcc tggagtcat tgccgttagc cagctccgcg 1200
gctccacca gggcaagatc ctctgcttct atggcccccc tggcgtgggt aagaccagca 1260
ttgctcgctc catcgccgc gccctgaacc gagagtactt ccgcttcagc gtcgggggca 1320
tgactgacgt ggctgagatc aagggccaca ggcggaccta cgtgggcgcc atgcccggga 1380
agatcatcca gtgtttgaag aagaccaaga cggagaaccc cctgatcctc atcgacgagg 1440
tggaacaagat cggccgaggc taccagggg acccgctcgtc ggcactgctg gagctgctgg 1500
accagagca gaatgccaac ttccctggacc actacctgga cgtgcccgtg gacttgtcca 1560
aggtgctgtt catctgcacg gccaacgtca cggacaccat ccccgagccg ctgcgagacc 1620

103

gtatggagat gatcaacgtg tcgggctacg tggcccagga gaagctggcc attgcggagc 1680
 gctacctgggt gccccaggct cgcgccctgt gtggcttggg tgagagcaag gccaagctgt 1740
 catcggacgt gctgacgctg ctcacaaagc agtactgccg cgagagcggg gtccgcaacc 1800
 tgcagaagca agtggagaag gtgttacgga aatcggccta caagattgtc agcggcgagg 1860
 ccgagtccgt ggagggtgacg cccgagaacc tgcaggactt cgtggggaag cccgtgttca 1920
 ccgtggagcg catgtatgac gtgacaccgc ccggcgtggg catggggctg gcctggaccg 1980
 caatgggagg ctccacgctg tttgtggaga catccctgag acggccacag gacaaggatg 2040
 ccaagggtga caaggatggc agcctggagg tgacaggcca gctgggggag gtgatgaagg 2100
 agagcgcccc catagcctac accttcgcca gagccttccct catgcagcac gccccgcca 2160
 atgactacct ggtgacctca cacatccacc tgcattgtgc cgaggcgcc accccaagg 2220
 acggcccaag cgcaggctgc accatcgta cggccctgct gtccctggcc atgggcaggc 2280
 ctgtccggca gaactctggc atgactggcg aagtctccct cacgggcaag atcctgcctg 2340
 ttggtggcat caaggagaag accattgcgg ccaagcgcg aggggtgacg tgcatcgtcc 2400
 tgccagccga gaacaagaag gacttctacg acctggcagc cttcatcacc gagggcctgg 2460
 aggtgcactt cgtggaacac taccgggaga tcttcgacat cgccttcccg gacgagcagg 2520
 cagaggcgct ggccgtggaa cggtagcggc caccocggga ctgcaggcgg cgatgtcag 2580
 gccctgtctg ggccagaact gagcgctgtg gggagcgcg ccggacctgg cagtggagcc 2640
 accgagcgag cagctcggtc cagtgacca gatcccaggg acctcagtcg gcttaatcag 2700
 agtgtggcat agaagctatt taatgattaa agtcatttgc agta 2744

<210> 60

<211> 2742

<212> DNA

<213> Homo sapien

<400> 60

cgaactccgc cccctcttct ccgcgtaggc ccagctccct gaagcggctg tttcgagcca 60
 cgcgcccata ggggtaccgag gcacgcgcgc ggcgtcacgt gcgtttcgcg gcgagcggaa 120
 atgacgcgag ttgtgtgagc cgccagtatg gccgggctat ggccgagagc actggctacg 180
 tgcgactgtg gggagcggcg cgggtgctggg tgctgcccgc gccgatgctg gccgcgcgcg 240
 gggggcgggt tcccactgca gcaggagcgt ggttgctccg aggccagcgg acctgcgacg 300
 cctctcctcc ttgggcaactg tggggccgag gcccggaat tgggggcaa tggcgggggt 360
 tttgggaagc gagcagccgc ggcggaggcg cattctcggg gggcgaggac gcctccgagg 420
 gcggcgcgga ggaaggagcc ggcggcgcg ggggcagcgc gggcgccggg gaaggcccgg 480

104

tcataacggc gctcacgccc atgacgatcc ccgatgtgtt tccgcacctg ccgctcatcg	540
ccatcacccg caaccogggtg ttcccgcgct ttatcaagat tatcgagggtt aaaaataaga	600
agttgggtga gctgctgaga aggaaagtgc gtctcgccca gccttatgtc ggcgtctttc	660
taaagagaga tgacagcaat gagtcggatg tggtcgagag cctggatgaa atctaccaca	720
cggggacgtt tgcccagatc catgagatgc aggacctgg ggacaagctg cgcgatgatcg	780
tcatgggaca cagaagagtc catatcagca gacagctgga ggtggagccc gaggagccgg	840
aggcggagaa caagcacaag ccccgagga agtcaaagcg gggcaagaag gaggcggagg	900
acgagctgag cgccaggcac ccggcggagc tggcgatgga gcccaccct gagctcccgg	960
ctgagggtgct catggtggag gtagagaacg ttgtccacga ggacttccag gtcacggagg	1020
aggtgaaagc cctgactgca gagatcgtga agaccatccg ggacatcatt gccttgaacc	1080
ctctctacag ggagtcagtg ctgcagatga tgcaggctgg ccagcgggtg gtggacaacc	1140
ccatctacct gagcgacatg ggcgcgcgc tcaccggggc cgagtcccat gagctgcagg	1200
acgtcctgga agagaccaat attcctaagc ggctgtacaa ggccctctcc ctgctgaaga	1260
aggaatttga actgagcaag ctgcagcagc gcctggggcg ggagggtggag gagaagatca	1320
agcagacca ccgtaagtac ctgctgcagg agcagctaaa gatcatcaag aaggagctgg	1380
gcctggagaa ggacgacaag gatgccatcg aggagaagtt ccgggagcgc ctgaaggagc	1440
tcgtggtccc caagcacgtc atggatgttg tggacgagga gctgagcaag ctgggcctgc	1500
tggacaacca ctctcggag ttcaatgtca cccgcaacta cctagactgg ctcacgtcca	1560
tcccttgggg caagtacagc aacgagaacc tggacctggc gcgggcacag gcagtgctgg	1620
aggaagacca ctacggcatg gaggacgtca agaaacgcat cctggagtgc attgccgtta	1680
gccagctccg cggctccacc cagggaaga tcctctgctt ctatggcccc cctggcgtgg	1740
gtaagaccag cattgctcgc tccatcgccc gcgccctgaa ccgagagtac ttccgcttca	1800
gcgtcggggg catgactgac gtggctgaga tcaagggccca caggcggacc tacgtgggcg	1860
ccatgcccgg gaagatcatc cagtgtttga agaagaccaa gacggagaa cccctgatcc	1920
tcatcgacga ggtggacaag atcgccgag gctaccaggg ggacctgctg tcggcactgc	1980
tggagctgct ggaccagag cagaatgcc acttcctgga ccactacctg gacgtgcccg	2040
tggactgtc caagggtgctg ttcatctgca cggccaacgt cacggacacc atccccgagc	2100
cgctgcgaga ccgtatggag atgatcaacg tgcggggcta cgtggcccag gagaagctgg	2160
ccattgcgga gcgtacctg gtgccccagg ctgcgcctt gtgtggcttg gatgagagca	2220
aggccaagct gtcacggac gtgctgacgc tgctcatcaa gcagtactgc cgcgagagcg	2280
gtgtccgcaa cctgcagaag caagtggaga aggtgttacg gaaatcggcc tacaagattg	2340

105

```

tcagcggcga ggccgagtcg gtggaggtga cggccgagaa cctgcaggac ttcgtgggga 2400
agcccggtgt caccgtggag cgcattgatg acgtgacacc gcccggcgtg gtcattgggc 2460
tggcctggac cgcaatgggt gagcgtgggt gcgggagaag accccaaagc cattccatt 2520
tctacccgag aacaagcagg tcccaccttg tacacctgtg cagtgggcct cagggtggctc 2580
tgaacggcct ctggcggggg gtggtgtgct gggagcccag ggggtctggg ctgagggagg 2640
ggaaggctgt gacacatggg ggtggcagca caccatgggc cttatggatt tgggtgccac 2700
tttgataggc ttacaggtaa ggacaagctc atccagatgg gc 2742

```

```

<210> 61
<211> 2116
<212> DNA
<213> Homo sapien

```

```

<400> 61
tgttgtagtg gctggagtac agtgggtgcag tctcagctca ctgcaacctc cgccctctcg 60
gttcaagcga ttctcctgcc tcagcctccc gagaagctgg aattacaggc atgcgccacc 120
acgcccagct aattttgtat ttttagtaaa gatgggtgtt ctccatgttg gtcaggctgg 180
tctcgaactc ccgacctcag gtgatccgcc cacctcggcc tcccaaagtg ctggcatgac 240
agagctagag ctggggccca gcccaggt gcagcccatt cgcaggcacc cgaaagaact 300
tcccagtat ggtggtcctg gaaaggacat ttttgaagat caactatctc ttctgtgca 360
ttccgatgga atttcagttc atcagatgtt caccatggcc accgcagaac accgaagtaa 420
ttccagcata gcggggaaga tgttgacca ggtggagaag aatcacgaaa aggagaagtc 480
acagcaccta gaaggcagcg cctcctcttc actctctctt gattagatga aactgttacc 540
ttaccctaaa cacagtattt ctttttaact tttttatttg taaactaata aaggtaatca 600
cagccaccaa cattccaagc taccctgggt acctttgtgc agtagaagct agtgagcatg 660
tgagcaagcg gtgtgcacac ggagactcat cgttataatt tactatctgc caagagtaga 720
aagaaaggct ggggataatt ggggtggctt ggttttgatt ttttgcttgt ttgtttgttt 780
tgtactaaaa cagtattatc ttttgaatat cgtagggaca taagtatata catgttatcc 840
aatcaagatg gctagaatgg tgcctttctg agtgtctaaa acttgacacc cctggtaaat 900
ctttcaacac acttccactg cctgcgtaat gaagttttga ttcattttta accactggaa 960
tttttcaatg ccgtcatttt cagttagatg attttgact ttgagattaa aatgccatgt 1020
ctatttgatt agtcttattt ttttattttt acaggcttat cagtctcact gttggctgtc 1080
attgtgacaa agtcaaataa accccaagg acgacacaca gtatggatca catattgttt 1140
gacattaagc ttttgccaga aaatgttgca tgtgttttac ctgacttgc taaaatcgat 1200

```

106

tagcagaaag gcatggctaa taatggtggg ggtgaaaata aataaataag taaacaaaat 1260
 gaagattgcc tgctctctct gtgcctagcc tcaaagcggt catcatacat cataccttta 1320
 agattgctat attttggggtt attttcttga caggagaaaa agatctaaag atctttttatt 1380
 ttcattctttt ttgggttttct tggcatgact aagaagctta aatggttgata aaatatgact 1440
 agttttgaat ttacaccaag aacttctcaa taaaagaaaa tcatgaatgc tccacaattt 1500
 caacatacca caagagaagt taatttctta acattgtggt ctatgattat ttgtaagacc 1560
 ttcaccaagt tctgatatct tttaaagaca tagttcaaaa ttgcttttga aaatctgtat 1620
 tcttgaaaaat atccttggtg tgtattaggt ttttaaatac cagctaaagg attacctcac 1680
 tgagtcatca gtaccctect attcagctcc ccaagatgat gtgtttttgc ttaccctaag 1740
 agagggttttc ttcttatttt tagataattc aagtgccttag ataaattatg ttttctttta 1800
 gtgtttatgg taaactcttt taaagaaaat ttaatatggt atagctgaat ctttttggtg 1860
 actttaaatc tttatcatag actctgtaca tatgttcaaa ttagctgctt gcctgatgtg 1920
 tgtatcatcg gtgggatgac agaacaaca tatttatgat catgaataat gtgctttgta 1980
 aaaagatttc aagttattag gaagcatact ctgtttttta atcatgtata atattccatg 2040
 atacctttat agaacaattc tggcttcagg aaagtctaga agcaatattt cttcaaataa 2100
 aagggtgtta aacttt 2116

<210> 62
 <211> 1855
 <212> DNA
 <213> Homo sapien

<400> 62
 gggggcgggg aagtggcggc tgaagcgccg ccggcggggc tcaactgtggt ggtgtgagtg 60
 ggaggcgggc gcgctggtgg ctgcagctgg ggtgaggcgc gaggcggcgc actcgacggc 120
 tgactggagc agcggtaaag gcgaggatgg agaccgaagg atataagtca aagagtacag 180
 cagaaaatgt ctactgaacg gacttcttgg acaagcctgt ccaccattca gaaaatagcc 240
 ctgggccttg ggatcccagc cagtgaaca gttgcctata tcctataccg caggtatagg 300
 gaaagcagag aagagcggct gacatttggt ggggaagatg acattgagat agagatgcgg 360
 gttccccagg aggctgtgaa actcatcatt ggccggcaag gagccaatat taaacagctg 420
 cggaaacaga caggtgctcg gattgatgtg gacacagagg atgtaggcga tgagcgagtg 480
 ctgcttatca gtggttttcc tgttcagggt tgcaaggcca aagcagcaat ccatcagatc 540
 ctgacagaga atacccaggt gtctgagcag ctttcagttc cccagagatc tgtgggcaga 600
 atcatagggg gaggcggcga gacaattcgt tctatctgta aggcattctg agccaaaatt 660

107

acctgtgaca aagaatcaga agggacatta ctactatcaa gacttataaa aatctcagga	720
acacagaagg aagtggcagc agccaagcat ttgatactgg agaaagtttc agaagatgaa	780
gaacttcgga agagaattgc tcattctgca gaaaccaggg tcccacgcaa acagccaatc	840
agtgtgagaa gagaagacat gacagagcca ggtggagctg gagagccagc attatggaaa	900
aacaccagtt ctagcatgga gccgactgca cccctgggtga ctccctccacc caaaggagga	960
ggcgacatgg ctgtggtagt gtcaaaggaa ggttcctggg agaaacctag tgatgacagc	1020
tttcagaagt ctgaagccca ggccatccca gagatgccca tgtttgaaag tatgtaacaa	1080
agagggagcc cattaatcat tgaagataga aatactggct tagatattgg agagattaga	1140
aggaatgcct tctcagtctg agcagccagt ataatatggc tggggtcagg acagaagatt	1200
cggtaagata ttttgaagat gggatggagg tagaatccac accaacagtt gctggacagt	1260
ttgaggatgc agatgttgat cactgaaaat gatttatgca ggtataagat tctgctccta	1320
attgtaggag agaacttggg gcctcttcca ctctgaagca aagttgatga aagtcttctt	1380
ccttttccca aaaccaacc tgaactactt ctttcttgga gacagactat attgagacaa	1440
gttgttacca gcaaaagata gatacatgac ctttattaac aaaaatgaat taaccaagag	1500
gatatttgta gtttattatt taccceaaac tttctgtgtc tgggtaccct ctgagtaggc	1560
ctataattcc tgccttcact gtatgcattt tatgtaagct agcagacctt tgtgggtgaga	1620
atgcacagga gcttggagggt ataaatagac aggggtgggaa agagagagct cctttcgcca	1680
tgttttacca gcccgctctg ttataacctc ttaggttata ttctttaatt tccaaccttt	1740
taggttagtt tctgtaacag aacaaatgag tctgggataa agtcctcaaa gtacttcaaa	1800
tggttaattgt tttgtttttg taacggotta acaataaacc taggttttct gttta	1855

<210> 63

<211> 1136

<212> DNA

<213> Homo sapien

<400> 63

sgctaggtac aggcggkgt kaatcgccg aggcgcacgt gctggaaggc ccagacacca	60
acaccacgat cattcagctg cagcccttgc aggagccga gagctggcg cgacgcaga	120
gtggcctgca gtcctacctg ctccagttcc acggcctcgt gcgcctggtg caccaggagc	180
ggaccttggc ctttctctg accatccgct gcttctctggg ctgtgagctg cctcccagg	240
gctctagagc ccatgtcttc ttcgaagtgg ctgtgaatgg gagctccttt gtgagtttcc	300
ggccggagag agccttgtgg caggcagaca cccaggtcac ctccggagtg gtcaccttca	360
ccctgcagca gctcaatgcc tacaaccgca ctcggtatga actgcgggaa ttcctggagg	420

108

```

acacctgtgt gcagtatgtg cagaaacata tttccgcgga aaacacgaaa gggagccaaa 480
caagccgctc ctacacttcg ctggtcctgg gcgtcctggg gggcagtttc atcattgctg 540
gtgtggctgt aggcattctc ctgtgcacag gtggacggcg atgttaatta ctctccagcc 600
ccgtcagaag gggctggatt gatggaggct ggcaaggga agtttcagct cactgtgaag 660
ccagactccc caactgaaac accagaaggt ttggagtgc agctcctttc ttctcccaca 720
tctgcccact gaagatttga gggaggggag atggagagga gaggtggaca aagtacttgg 780
tttgctaaga acctaagaac gtgtatgctt tgctgaatta gtctgataag tgaatgttta 840
tctatctttg tggaaaacag ataatggagt tggggcagga agcctatggc ccctcctcca 900
aagacagaca gaatcacctg aggcgttcaa aagatataac caaataaaca agtcatccac 960
aatcaaaata caacattcaa tacttccagg tgtgtcagac ttgggatggg acgctgatat 1020
aatagggtag aaagaagtaa cacgaagaag tgggtggaaat gtaaaatcca agtcatatgg 1080
cagtgatcaa ttattaatca attaataata ttaataaatt tcttatattt aaggca 1136

```

```

<210> 64
<211> 2320
<212> DNA
<213> Homo sapien

```

```

<400> 64
cgggatagct gtcttgagcc ccaagcctct tctcccctg ctgcccctct gcagccattc 60
gggatgggac cccctctggg gtgtcagcac gaaagggcta acgggagccc ctctccttggc 120
ctccccctgt aggttacaga gccatcacgg tgcggaagct catgcagggc atgggccttg 180
gcctctccag cgtctttgct ctgtgcctgg gccacacctc cagcttctgt gagtctgtgg 240
tctttgcata agcctccatc ggctccaga ccttcaacca cagtggcatt tctgttaaca 300
tccaggactt ggccccgtcc tgcgccggtt ttctgtttgg tgtggccaac acagccgggg 360
ccttggcagg tgtcgtgggt gtgtgtctag gcggctactt gatggagacc acgggctcct 420
ggacttgctt gttcaacctt gtggccatca tcagcaacct ggggctgtgc accttctgg 480
tgtttggaca ggctcagagg gtggacctga gctctacca tgaggacctc tagctcccaa 540
ccccacagcc tctccaagga ccagggcgc agcagccccg ggacacaggg gactcagtgt 600
gtgggacttg gtcactccat gtcagacaca cgagcagaga ggaacacaaa ccactgtgga 660
gcctgaagct ccttaagaag agtccacaac agctggtggg aggggtggggg gggcctgggt 720
ccagaccagg ctgcgtgctc tctgggcctc agtttcccca cctgccagcg ggctcgggcc 780
tgtctcctc acaggctggg gtggccgtca ggggtgggtg gggtattgtt agtaggcgca 840
gcctcattcc caccacgata tgttccgctt gggtcccgcc aaacctccct cggctcgccgt 900

```

gttctccgca agcctcctgc agcgcccgcc tgccaatgtg aggctggcac caggctgcag 960
 cctccccaat ccagagccac tttgctgtgt ctctggcggg ctgtcctcct tgggtgggagc 1020
 tgtcctgcac actgtaggat gcttaaaggt atccctggcc tccacccatc cctagccagc 1080
 agctccagct cagacaacag ccagaaatgt ctccagactc tgcccagcct cccagggtag 1140
 ccaccctcga gacatgacct cagagtctct gtgtctccta gaagcctgac agagaccccc 1200
 agggcagtggt gtgggtggcg ggctagagac ccttgccctgt gtccgggacc ctggcgccgc 1260
 tctccctcc tgtggatccc tccgcactaa cagtgttctc agtgggcaga cgcctgggca 1320
 ccccttgggc cctgcccagc atggccatgg cgcaggctct cgaaccgcga tggctttccc 1380
 aggcctgggtg attctgctct ccagggaagg ttggcacctt cctcgggggc gggccccacg 1440
 cccccagaa cacacagacc cacctttctg gcgttcttcc tacctccctt ttcgttgctt 1500
 gaggagctgg tggtttcatg agttaatgat acatcttgca aggtgtacac atagagaaaa 1560
 aaacctaaaa atgtggaaaa gcacgcaaaa gccttattta aataataact attaaactat 1620
 tcaaaaagaa aaaaaaaaaa caaacaaaaa acataaaaac aaaaaaaca aaagaaaaaa 1680
 aaaaaaaca aaacaggggg gggggcgcg cccagaagaa atcccgcccg cggggggcg 1740
 cccaataac cgcgcgacca cccggcttcc ttgttagcac acaacagggg gggggcgccc 1800
 acaagaacga gcgagcgac cgtagtcata actatacgca gcgcgagctc cgatggccgc 1860
 gccacacac ctcaaccaca cagtcgtggg tacaccctgc ggcaaccaa cgctggccac 1920
 aaccggcggtg aaacatcttg cgtggagaga aaggcgacc gctcacacta cccagcgag 1980
 agcgagggac ggcaaccaa acatcggtgt gggggccgc gacaaacctc cctcactcct 2040
 caaccatgcg cgacgatctt cattacaaa tcgccccct cgctcgacgg gggcgacga 2100
 atcgctcatc acacacaatc catcttattc acacggcgt cgcgctatcc tccacgcgtg 2160
 tgtggtgcct catcacaccg aacgacgggt cgcataata atatgtcgac cgggccgcg 2220
 ctagctcgtc aggaaaatga agatttcgtg ctggctcaat cttaggaggc agaaattgca 2280
 ctcgatcggt cgtcgctgac gaacacacat ttctttctcc 2320

<210> 65
 <211> 1609
 <212> DNA
 <213> Homo sapien

<400> 65
 agcggggaga aacgttgag cccgcgccga acgcccggca gcacaaagga tccccgactg 60
 ccggggagcg gtgctcgag ggcacaggtc tacgccatcc cccacgcagt ttcggagatg 120
 gagcgctggg cccatggagg gaaggcgga ggctcgggg ctccggcagc ttgctggggc 180

110

```

aggggctcaa ggcggcagtc cgatagtggg ggccgctgag aactgtcacg gagctgcgtc      240
tgtacagcga gcatccctta tttattcagg gcgagtgtgt atttggggcg gcgtgcaggg      300
ggctgacaaa gaccggagag ctcccgggtgc ggccgccggc ggagcgaaga ctggaacccg      360
tatgagcgcc cccagcgccc cctgagcgct cgccgccggg gcacggcgca ccccgcgga      420
ggcaggggatc agcaaagccg tgcgccccga ggcccgcccc cgtctccgca caaagaccga      480
gctggaggat cttcagaaga agcctcccc atacctgcgg aacctgtcca gcgatgatgc      540
caatgtcctg gtgtggcacg ctctcctcct acccgaccaa cctccctacc acctgaaagc      600
cttcaacctg cgcacagct tcccgccgga gtatccgttc aagcctccca tgatcaaatt      660
cacaaccaag atctaccacc ccaacgtgga cgagaacgga cagatttgcc tgcccatcat      720
cagcagtgag aactggaagc cttgcaccaa gacttgccaa gtcttgagg ccctcaatgt      780
gctggtgaat agaccgaata tcagggagcc cctgcggatg gacctcgctg acctgctgac      840
acagaatccg gagctgttca gaaagaatgc cgaagagttc accctccgat tcggagtgga      900
ccggccctcc taactcatgt tctgaccctc tgtgcaactgg atcctcggca tagcggacgg      960
acacacctca tggactgagg ccagagcccc ctgtggccca ttccccattc atttttccct     1020
tcttaggttg ttagtcatta gtttgtgtgt gtgtgtggtg gagggaaggg agctatgagt     1080
gtgtgtgttg tgtatggact cactcccagg ttcacctggc cacagggtga cccttcccac     1140
accttttaca ttccccagag ccaagggagt ttaagtttgc agttacaggc cagttctcca     1200
gctctccatc ttagagagac aggtcacctt gcaggcctgc ttgcaggaaa tgaatccagc     1260
agccaactcg aatcccccta gggctcaggc actgagggcc tggggacagt ggagcatatg     1320
gggtgggagac agatggaggg taccctatctt acaactgagt cagccaagcc actgatggga     1380
atatacagat ttaggtgcta aaccgtttat tttccacgga tgagtcacaa tctgaagaat     1440
caaacttcca tcctgaaaat ctatatgttt caaaaccact tgccatcctg ttagattgcc     1500
agttcctggg accaggcctc agactgtgaa gtatatatcc tccagcattc agtccagggg     1560
gagccacgga aacctgttcc ttgcttaagc cattaaagtc agagatgaa                     1609

```

```

<210> 66
<211> 1414
<212> DNA
<213> Homo sapien

```

```

<400> 66
tcctgcctga aacttgtgat aacaactggg tgcattctca gtgcaggaaa agtccttggc      60
atcttggtgc caggcattct gctactgctg gtggtgggtt atgaacacag atgaccagga     120
tgaagtcctt gccttggaat aagcccatag cccatcccat aatttgatgc ctgtattact     180

```


111

```

agagtcctga gaaacacgga atgctgatta gaatttggga caaagttccc tgagtattct      240
ctggaggaaa agatgctggag gtggacctgg aaggcctcct ggagagctgg aggatcttca      300
gaagaagcct ccccatatac tgcggaacct gtccagcgat gatgccaatg tccctgggtgtg      360
gcacgctctc ctctaccctg accaacctcc ctaccacctg aaagccttca acctgcgcac      420
cagcttcccg ccggagtatc cgttcaagcc tcccatgatc aaattcacaa ccaagatcta      480
ccaccccaac gtggacgaga acggacagat ttgcctgccc atcatcagca gtgagaactg      540
gaagccttgc accaagactt gccaaagcct ggaggccctc aatgtgctgg tgaatagacc      600
gaatatcagg gagcccctgc ggatggacct cgctgacctg ctgacacaga atccggagct      660
gttcagaaaag aatgccgaag agttcacctc ccgattcgga gtggaccggc cctcctaact      720
catgttctga ccctctgtgc actggatcct cggcatagcg gacggacaca cctcatggac      780
tgaggccaga gcccctgtg gccattccc cattcatttt tcccttctta ggttgttagt      840
cattagtttg tgtgtgtgtg tgggtggaggg aaggaggcta tgagtgtgtg tgttgtgtat      900
ggactcactc ccagggtcac ctggccacag gtgcacctt cccacacctt ttacattccc      960
cagagccaag ggagtttaag tttgcagtta caggccagtt ctccagctct ccatcttaga     1020
gagacaggtc accttgacgg cctgcttgca ggaaatgaat ccagcagcca actcgaatcc     1080
ccctagggct caggcactga gggcctgggg acagtggagc atatgggtgg gagacagatg     1140
gagggtaccc tatttacaac tgagtcagcc aagccactga tgggaatata cagatttagg     1200
tgctaaaccg tttattttcc acggatgagt cacaatctga agaatcaaac ttccatcctg     1260
aaaatctata tgtttcaaaa ccacttgcca tctgttaga ttgccagttc ctgggaccag     1320
gcctcagact gtgaagtata taccctccag cattcagtc ccagggggagcc acggaaacca     1380
tgttcttgct taagccatta aagtcagaga tgaa                                     1414

```

<210> 67

<211> 1243

<212> DNA

<213> Homo sapien

<400> 67

```

agcgggggaga aacgttgacg cccgcgccga acgccgggca gcacaaagga tccccgactg      60
ccgggggagcg gtgctcggag ggcacaggtc tacgccatcc cccacgcagt ttcggagatg     120
gagcgcctggg cccatggagg gaaggcggca ggggaacctg tccagcgatg atgccaatgt     180
cctgggtgtgg cacgctctcc tctaccctga ccaacctccc taccacctga aagccttcaa     240
cctgcgcacac agcttcccgc cggagtatcc gttcaagcct cccatgatca aattcacaaac     300
caagatctac caccccaacg tggacgagaa cggacagatt tgctgcccac tcatcagcag     360

```

112

tgagaactgg aagccttgca ccaagacttg ccaagtctg gaggccctca atgtgctggt 420
 gaatagaccg aatatcaggg agcccctgcg gatggacctc gctgacctgc tgacacagaa 480
 tccggagctg ttcagaaaga atgccgaaga gttcaccctc cgattcggag tggaccggcc 540
 ctccctaactc atgtttctgac cctctgtgca ctggatcctc ggcatagcgg acggacacac 600
 ctcattggact gagggccagag ccccctgtgg ccattcccc attcattttt cccttcttag 660
 gttgttagtc attagtttgt gtgtgtgtgt ggtggaggga agggagctat gagtgtgtgt 720
 gttgtgtatg gactcactcc caggttcacc tggccacagg tgcacccttc ccacaccctt 780
 tacattcccc agagccaagg gagtttaagt ttgcagttac aggccagttc tccagctctc 840
 catcttagag agacaggtca ccttgccaggc ctgcttgag gaaatgaatc cagcagccaa 900
 ctccaatccc cctagggctc aggcactgag ggcctgggga cagtggagca tatgggtggg 960
 agacagatgg aggggtaccct atttacaact gagtccagca agccactgat gggaatatac 1020
 agatttaggt gctaaaccgt ttattttcca cggatgagtc acaatctgaa gaatcaaact 1080
 tccatcctga aaatctatat gtttcaaac cacttgccat cctgtagat tgccagttcc 1140
 tgggaccagg cctcagactg tgaagtatat atcctccagc attcagttca gggggagcca 1200
 cggaaccat gttcttgctt aagccattaa agtcagagat gaa 1243

<210> 68
 <211> 1507
 <212> DNA
 <213> Homo sapien

<400> 68
 ggggggtggg tccccggggc ggggcggggc gcgctgtgtc gcgggtcgga gctcggctcct 60
 gctggaggcc acgggtgcca cacactcggc cccgacatga tggcgagcat gcgagtgggtg 120
 aaggagctgg aggatcttca gaagaagcct ccccatatcc tgcggaaact gtccagcgat 180
 gatgccaatg tcctgggtgtg gcacgctctc ctctaccgc aggcagaggt tgcgggtgagc 240
 cgagatcacg ccattgcact ccagcctggg cagcaaagt aaactccatc tcaaaaaaaaa 300
 aaaaaaaaaa agccttgcca ccagcatctc cttctgatga ggcctcagg aagcttccac 360
 tcatgggtgga aggcaaagg gagccacgtg tacagatcac attgcaagag aagaagtgaa 420
 agagagagag agcgagcagg tgccaggatc ttctaaacaa ccagcttttt cagaccaacc 480
 tccctaccac ctgaaagcct tcaacctgcg catcagcttc ccgccggagt atccgttcaa 540
 gcctcccatg atcaaatca caaccaagat ctaccacccc aacgtggacg agaacggaca 600
 gatttgcttg cccatcatca gcagtgagaa ctggaagcct tgcaccaaga cttgccaagt 660
 cctggaggcc ctcaatgtgc tggatgaatg accgaatatc agggagcccc tgcggatgga 720

113

```

cctcgctgac ctgctgacac agaatccgga gctgttcaga aagaatgccg aagagttcac      780
cctccgattc ggagtggacc ggccctccta actcatgttc tgaccctctg tgcactggat      840
cctcggcata gcggacggac acacctcatg gactgaggcc agagccccct gtggcccatt      900
ccccattcat ttttcccttc ttaggttggt agtcattagt ttgtgtgtgt gtgtgggtgga      960
gggaagggag ctatgagtgt gtgtgttggt tatggactca ctcccagggt cacctggcca     1020
caggtgcacc cttccacac cctttacatt ccccagagcc aaggaggttt aagtttgcag     1080
ttacaggcca gttctccagc tctccatctt agagagacag gtcaccttgc aggctgctt     1140
gcaggaaatg aatccagcag ccaactcgaa tccccctagg gtcaggcac tgagggcctg     1200
gggacagtgg agcatatggg tgggagacag atggagggta ccctatttac aactgagtca     1260
gccaagccac tgatgggaat atacagattt aggtgctaaa ccgtttattt tccacggatg     1320
agtcacaatc tgaagaatca aacttccatc ctgaaaatct atatgtttca aaaccacttg     1380
ccatcctggt agattgccag ttcctgggac caggcctcag actgtgaagt atatatcctc     1440
cagcattcag tccaggggga gccacggaaa ccatgttctt gcttaagcca ttaaagtcag     1500
agatgaa                                           1507

```

```

<210> 69
<211> 1279
<212> DNA
<213> Homo sapien

```

```

<400> 69
cagcggcccg ccgcggggcc tttttgtcct gagggccaga gaaatggaga aggggggtggg      60
gggacagcca cgtggccgca ggaggattta caacattttc tttcgccatc gatgttatcg     120
caaaatgtgt gagagaagcg gctgcgcagc ccggacggga gcgtgagggg gcggggccagg     180
taagcagccc cggcggtttc gccgcatacg ggactgcggg gcgaccgcgg gcaccagcca     240
cgcgacgcgg ctccgcgggg tctcggccgg gtccgcgctc tgaaggatct cgagagccat     300
ggatgggtgca ggcgggaccc tcgagctgca gcatctccgg tgaccggggg ttgccgagga     360
gggtggagacc agcacagggt gtccggcccc ggcgccctccg aatccggggg tgggtcaagac     420
ggatccccaa ggctgaggtc ggcagtcccc gggactcgca gctgttgagc ctgtggagac     480
gcggccccgt gaccgaggca cccttcagca acccgggggc agcgttccag agactgaact     540
tctcaaataca ctgcttcaac agcttttaaa aatctggacc tagttactcc tgtcatctat     600
gtgtaaagat ttagaaaaaa aatcccaaac ccaaggggtg gcagccgcga ggatgtacag     660
ctcctggcag tgtgcgctca cccggagacc tgagggtcaaa gcccaaatgc tgagcacctc     720
ctgggaggcg ggagcatggt ctgctcagtg ccagacatct ggatattgag aaaacatctg     780

```

114

```

.ggtagctggg gttttcaggt ttctgcctga catttaatat cacaaacatt gagcctgctc      840
ctgtccccc cagccacgc ggcttctcct cccaaacata ccagctcct tcccctgcct      900
tcctgacggg taagaaacat ccattctttc caattcccca gcgttttccc cctaccggaa     960
atctgatggg cttatgacat catggctggc tgctgagcga tgaagtggat gccacaaaga    1020
aatccgacat atcagataga ttctgaaatc ggtttccctc cagctgtagt aacaggcgtg    1080
aagtcaggag aatttgagct ttgtttaaaa aataaataaa taaataaata aaccataaca    1140
aagtccttggc ctgtattaaa tgcaattttc ttaaaaacaa gcaaaccctt tggacatcat    1200
tttattttta tagaaatgct gagttttatg aaactaaagt ggctaataaa tcagacctga    1260
agctttgtgt gagtgttcc                                     1279

```

```

<210> 70
<211> 517
<212> DNA
<213> Homo sapien

```

```

<400> 70
cagcggccgg ccgcggggcc tttttgtcct gagggccaga gaaatggaga aggggggtggg      60
gggacagcca cgtggccgca ggaggattta caacattttc ttctgccatc gatgttatcg     120
caaaatgtgt gagagaagcg gctgcgcagc ccggacggga gcgtgagggg gcggggccagc     180
gttttcccc taccggaaat ctgatgggct tatgacatca tggctggctg ctgagcgatg     240
aagtggatgc cacaaagaaa tccgacatat cagatagatt ctgaaatcgg tttccctcca     300
gctgtagtaa caggcgtgaa gtcaggagaa tttagctttt gtttaaaaaa taaataaata     360
aataaataaa ccataacaaa gtcttgcctt gtattaaatg caattttctt aaaaacaagc     420
aaaccttttg gacatcattt tattttaata gaaatgctga gttttatgaa actaaagtgg     480
ctaataaatc agacctgaag ctttgtgtga gtgttcc                                     517

```

```

<210> 71
<211> 736
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (46)..(46)
<223> n=a, c, g or t

```

```

<400> 71
gctcccaaaa gggtagctgc ttctgctggc ttaatgcctc agaacnttgg tgtcattggg      60
ctcagacacc actttgccat ccactatccg gcgggtgggt gtctttggga tggtttgcac    120

```

115

```

ggagttgctg ctgtccaagg catcaccaag attaaagtcc tcgccatcta cagctttctgc 180
tcccagatct gtgacccccg gacgaccggc gcgttctggc agacgtggaa ggactttgag 240
gtccggcatg gcaatgagga caccatcaag gaaatgctgc gtatccggcg cagcgtgcag 300
gccacgtaca acacgcaggt caacttcatg gcctcgcaga tgctcaaggt ctcgggcagt 360
gccacgggca ccgtgtctga cctggccctt gggcagagtg gcatggacga catgaagctg 420
ctggaacagc gggcagagca gctggcggct gaggcggagc gtgaccagcc cttgcgcgcc 480
cagagcaaga tcctgttcgt gaggagtgc gcctcccgga aggagctggc agagctggca 540
cagcaggtca accccgagga gatccagctg ggcgaggacg aggacgagga cgagatggac 600
ctggagccca acgaggttcg gctggagcag cagagcgtgc cagccgcagt gtttgggagc 660
ctgaaggaag actgaccctt cctccccc cccccctcc ccacccctc cccaatacag 720
ctacgtttgt acatca 736

```

```

<210> 72
<211> 3041
<212> DNA
<213> Homo sapien

```

```

<400> 72
aatgatagat cactataggg gaatgggtcc tetaatgctg ctcgagcggc gcagtgtgat 60
ggatcgtggt cgcggcgagg taccagtatt ctccacagaa ccagcatgta gagcagcagc 120
cacactacac ccacaaacca actctggaat acagtccttt tcccatacct cccagctccc 180
ccgcttatga accaaacctc tttgatggtc cagaatcaca gttttgcccc aaccaaagct 240
tagtttccct tcttggtgat caaagggaat ctgagaatat tgctaataccc atgcagactt 300
cctccagtgt tcagcagcaa aatgatgctc acttgcacag cttcagcatg atgccagca 360
gcgcctgtga ggccatggtg gggcacgaga tggcctctga ctcttcaaac acttcaactgc 420
cattctcaaa catgggaaat ccaatgaaca ccacacagtt agggaaatca ctttttcagt 480
ggcaggtgga gcaggaagaa agcaaattgg caaatatttc ccaagaccag tttctttcaa 540
aggatgcaga tggtgacacg ttccttcata ttgctgttgc ccaagggaga agggcacttt 600
cctatgttct tgcaagaaag atgaatgcac ttcacatgct ggatattaaa gagcacaatg 660
gacagagtgc ctttcaggtg gcagtggctg ccaatcagca tctcattgtg caggatctgg 720
tgaacatcgg ggcacaggtg aacaccacag actgctgggg aagaacacct ctgcatgtgt 780
gtgctgagaa gggccactcc caggtgcttc aggcgattca gaaggagca gtgggaagta 840
atcagtttgt ggatcttgag gcaactaact atgatggcct gactcccctt cactgtgcag 900
tcatagccca caatgctgtg gtccatgaac tccagagaaa tcaacagcct cattcacctg 960

```

116

aagttcagga gcttttactg aagaataaga gtctggttga taccattaag tgcctaattc 1020
aatggggagc agcgggtggaa gcgaaggatc gcaaaagtgg ccgcacagcc ctgcatttgg 1080
cagctgaaga agcaaatctg gaactcattc gcctcttttt ggagctgccc agttgcctgt 1140
cttttgtgaa tgcaaaggct tacaatggca aactgcccct ccatgttgct gccagcctgc 1200
agtatcgggt gacacaatta gatgctgtcc gcctgttgat gaggaaggga gcagacccaa 1260
gtactcggaa cttggagaac gaacagccag tgcatttggg tcccgatggc cctgtgggag 1320
aacagatccg acgtatcctg aagggaagt ccattcagca gagagctcca ccgtattagc 1380
tccattagct tggagcctgg ctagcaacac tctctgtcag ttaggcagtc ctgatgtatc 1440
tgtacataga ccatttgcct tatattggca aatgtaagtt gtttctatga aacaaacata 1500
tttagttcac tattatatag tgggttatat taaaagaaaa gaagaaaaat atctaatttc 1560
tcttggcaga tttgcatatt tcataccag gtatctggga tctagacatc tgaatttgat 1620
ctcaatggta acattgcctt caattaacag tagcttttga gtaggaaagg actttgattt 1680
gtggcacaaa acattattaa tatagctatt gacagtttca aagcaggtaa attgtaaatg 1740
tttctttaag aaaaagcatg tgaaaggaaa aaggtaata cagcattgag gcttcatttg 1800
gccttagtcc ctgggagtta ctggcgttgg acaggcttca gtcattggac tagatgaaag 1860
gtgtccatgg ttagaatttg atctttgcaa actgtatata attgttattt ttgtccttaa 1920
aaatattgta catacttggg tgtaacatg gtcataattg aaatgtataa gtccataaaa 1980
tagaaaagaa caagtgaatt gttgctattt aaaaaattt tacaattctt actaaggagt 2040
ttttattgtg taatcactaa gtctttgtag ataaagcaga tggggagtta cggagttgtt 2100
cctttactgg ctgaaagata tattcgaatt gtaaagatgc tttttctcat gcattgaaat 2160
tatacattat ttgtagggaa ttgcatgctt tttttttttt ttctcccgag acagggtctt 2220
gctctggcgc ccaggetgga gtacagtggc atgatcttgg ctcacttcag ccttgacttg 2280
ggctcaagtg atcctcctac ctgagccttc tgagtaactg ggactacagg tgtgcactcc 2340
tcgcctggct aattttttat tttttgtaca ggcaggatct tgccaccttg ccaggetgg 2400
tcttgaactc ctgagctcat gccatctgcc tgccttagtc tcccaaatg ctgggattac 2460
aggagtgagc caccatgccc ggctggcagt tgcattggaag agaacacctc tttatggctt 2520
accctctaga atttctaatt tatgtgttct gttgaaattt ttgttttttt acctttattg 2580
aaacaacaaa aagtcagtat tgaaacatat cttcctgttt tctgttgtca aatgatgata 2640
atgtgccatg atgttttata tatacattc agaaaaagt ttatttttta ataacattct 2700
attaacatta ttttgcttgc cgctggcatg cctgaggaat gtatttggct ttgattacac 2760
actaagtttt tgtaataaat ttgactcatt aaaaaccttt tttttttaa aaaaaaaaaa 2820

117

agaaaatctc attagtgaac ttatctttgc agctgagtac tttaaattctt tttaaaaaga 2880
 tacccttttg attgatcaca ttgtttgacc cagtatgtct tgtagacacg ttagttataa 2940
 tcaccttgta tctctaaata tgggtgtgata tgaaccagtc cattcacatt ggaaaaactg 3000
 atgggttttaa ataaactaat tcactaatat tatttgtctt a 3041

<210> 73
 <211> 1193
 <212> DNA
 <213> Homo sapien

<400> 73
 acatccagct gcctgagacc ctctgcagc cttctcaagg gacagcccca ctctgcctct 60
 tgctcctcca gggcagcacc atgcagcccc tgtggctctg ctgggcactc tgggtgttgc 120
 ccctggccag ccccggggcc gccctgaccg gggagcagct cctgggcagc ctgctgcggc 180
 agctgcagct caaagaggtg cccaccctgg acagggccga catggaggag ctggtcatcc 240
 ccacccacgt gagggcccag tacgtggccc tgctgcagcg cagccacggg gaccgctccc 300
 gcggaaagag gttcagccag agcttccgag aggtggccgg cagggttcctg gcgttgagg 360
 ccagcacaca cctgctggtg ttccggcatgg agcagcggct gccgcccaac agcgagctgg 420
 tgcaggccgt gctgcggctc ttccaggaat gtacattgac ctgcagggga tgaagtgggc 480
 cgagaactgg gtgctggagc ccccgggctt cctggcttat gagtgtgtgg gcacctgccg 540
 gcagcccccg gaggccttg ccttcaagtg gccgtttctg gggcctcgac agtgcacgc 600
 ctcgagact gactcgctgc ccatgatcgt cagcatcaag gagggaggca ggaccaggcc 660
 ccagggtggtc agcctgcccc acatgagggg gcagaagtgc agctgtgcct cggatggtgc 720
 gctcgtgcc aaggagctcc agccataggc gcctagtgtg gccatcgagg gacttgactt 780
 gtgtgtgttt ctgaagtgtt cgagggtacc aggagagctg gcgatgactg aactgctgat 840
 ggacaaatgc tctgtgctct ctagttagcc ctgaatttgc ttctctgac aagttacctc 900
 acctaatttt tgcttctcag gaatgagaat ctttggccac tggagagccc ttgctcagtt 960
 ttctctattc ttattattca ctgcactata ttctaagcac ttacatgtgg agatactgta 1020
 acctgagggc agaaagccca atgtgtcatt gtttacttgt cctgtcactg gatctgggct 1080
 aaagtcctcc accaccactc tggacctaa acctgggggt aagtgtgggt tgtgcatccc 1140
 caatccagat aataaagact ttgtaaaaca tgaataaaac acattttatt cta 1193

<210> 74
 <211> 1725
 <212> DNA
 <213> Homo sapien

118

<400> 74
tttttttttca cctaaattgt ttttacttgc ttcaccttaa atcagttatt cggttccttt 60
gaagaacatg tagcttttagc aacccccaaa gtgaaccagc ataccagaa ctactgattg 120
ttacatcaga gtttttgaaa acgccaatac tggaaacagg cgcacagtta aggagaggca 180
ggcctcaaac taaagcagaa ctttctcaaa actacttata tttcaagaga tggattaagc 240
gaaaagagct gtacaaagaa ataagccatt ttctatgcgc ctgttatatg caaggcactg 300
cactcccttg tcaacttgta atttacaaga agactcgccg gtccagggaa gggagagggg 360
aatagggtgca aggtttgtgg agggcggcaa actgcactcc atgtggcgcg aggctgcgct 420
tgcggagatc cgggttcttg tgcctggcgc cggcagccag taggggcgcg cgtcctgcgc 480
agttgccgcg cgcggcctcc tcgggcttca ggccgggcca atcaggaggc gggcgggagg 540
gggcgcggta ccaggaggga acgggacggg tcgggcttcg aggtcagctg gtccgcaggg 600
aagccttttg cttccccacc ggcaatggag cggttgacgt tgcctctcgg cggcgcggcg 660
gcggtggacg agtacctgga gtaccggagg tataaacaac ataaaactga cttggaagcg 720
attcctcagc agtgcccat tgatctgccc tgccaagtga ctggctgcca gtgcagggct 780
tacctttatg tccccttgaa tggtagccag cccattcgct gcaggtgcaa acactttgct 840
gatcagcaca gtgctgcgc tggctttaca tgcaatacat gttccaagt ttcaggattc 900
catagctgct tcaacttgtc ttgtggtcag cctgcatatg cccatgacac agtagtgga 960
actaagcaag aaagattggc tcaggaaaaa ccagtgggac aggacattcc ttatgcagcc 1020
atgggaggat taactggttt cagctcgctg gcggaaggct acatgcggtt agatgacagt 1080
gggattgtag gtacaagtag tcaagtttct tcattaagga gacctgaaga ggatgatatg 1140
gctttctttg aaagacgata ccaggaaagg atgaaaatgg aaaaggctgc taagtggaaa 1200
ggaaaagctc cattgccatc agctacaaaa cttcatgaa gactattgga gaaattaaaa 1260
ccatcatcca agtatctttt tcatgtttat ttaaatgtaa taatacagtt tatttttcct 1320
gaaattatct actttttttt tttactgtat aaatgtcttt tgggatgttt ccttaattta 1380
tttaaataac taaaaatgtc tattactttt gtcaaaactc ataatttact actttgtatg 1440
tacctttctt tctcctgaca aatgagggtta ttttatatga gtctgtctga gagtacagta 1500
aatgttttta gtacataata atttaactgt ttcagggtatt taaaaatta aagatattat 1560
caagggtttt ggacaaacat atgagccatt tttttgtcat tcaaaatagt aagttaaaaa 1620
caagagaaca aaagaaaaaa atagttaaaa tcattaattt ttttattttt caaactttgt 1680
aatgttatgt tttcaaataa aaactatctc aaaattttac aacca 1725

119

<210> 75
 <211> 670
 <212> DNA
 <213> Homo sapien

<400> 75
 ggagggttcag gaagccattt tctttcgcgt ctgcggtgct cggagtgtgg tacttctcct 60
 agttgcagtc aggcttcata cgctattgtc ctgcccgtta gagcagccag cgggtacaga 120
 atggattttg gaagagggag tcaccactgg acctccaagg aagccacgtg cagacatcta 180
 caaccttcga tctcctgacg agtttattgt tggccaaaac caggctttga ttgaaccagg 240
 atgaatgcgg gtgttggaag tagaatatat atatacatat aaaattgggtt gggagccacg 300
 tgtaccagtg tgtgttgatc ttggcttgat tcagtctgcc ttgtaacaga aactggcgat 360
 ggaatatgag aggagccctc tggaaagaaa aggacagacc ctgtgctttc atgaaagtga 420
 agatctggct gaaccagttc cacaaggtta ctgtatacat agcctgagtt taaaaggctg 480
 tgcccacttc aagaatgtca ttgttagact ttgaaatttc taactgccta cctgcataaa 540
 gaaaataaaa tcttttaaat caaaaaaaaa acaaaaacaa aaaaacttgt ggtggggggc 600
 cttggggccc ggaaaagggt ttttaaccct tccgttgggg ggggggggcc ccttggtttg 660
 cggccccctg 670

<210> 76
 <211> 964
 <212> DNA
 <213> Homo sapien

<400> 76
 ggagggttcag gaagccattt tctttcgcgt ctgcggtgct cggagtgtgg tacttctcct 60
 agttgcagtc aggcttcata cgctattgtc ctgcccgtta gagcagccag cgggtacaga 120
 atggattttg gaagagggag tcaccactgg acctccaagg aagccacgtg cagacatcta 180
 caaccttcga tctcctgacg agtttattgt tggccaaaac caggctttga ttgaaccagg 240
 atgaatgcgg gtgttggaag tagaatatat atatacatat aaaattgaaa ctggcgatgg 300
 aatatgagag gagccctctg gaaagaaaag gacagaccct gtgctttcat gaaagtgaag 360
 atctggctga accagttcca caaggttact gtatacatag cctgagttta aaaggctgtg 420
 cccacttcaa gaatgtcatt gttagacttt gaaatttcta actgcctacc tgcataaaga 480
 aaataaaatc ttttaaatca aaaaaaaaa aaaggggggg ccgccctttt tttttttttt 540
 tttgggaggt cccaacaat tcttgggtaa ttctttcccc aaaaggtaaa tatacccgaa 600
 ggggtttttta agttccacaa aaaaaacctt atttgggttc cttcccaaaa gaaaaaaaa 660
 atacatgaga gcaacagccc agcccgatag acagcacaga gacaactcgc aaccaacaag 720

120

```

tccgccaac gacaagcaga aacaaaagga gggcggcgag ccaacacatc agcaaggcga      780
caagaccac agccaagagg caggcggccg acagtgcacat aacaacgagc ggccccacgg      840
caacgacaac cgacgataag aacgacgtgt gtgaagacac accccaccgc cgaacgacgg      900
gatggcacca gcgagaccta tagtacgcga agcagccaag acgaagagcg gggaagcaag      960
gaaa                                                                    964

```

```

<210> 77
<211> 269
<212> DNA
<213> Homo sapien

```

```

<400> 77
cctctctggc ggcggctgct gctagggagt cgacttctcc tgccttgtaa cagaaactgg      60
cgatggaata tgagaggagc cctctggaaa gaaaaggaca gaccctgtgc tttcatgaaa      120
gtgaagatct ggctgaacca gttccacaag gttactgtat acatagcctg agttttaaag      180
gctgtgcca cttcaagaat gtcattgtta gactttgaaa tttctaactg cctacctgca      240
taaagaaaat aaaatctttt aaatcaaaa                                     269

```

```

<210> 78
<211> 928
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (855)..(855)
<223> n=a, c, g or t

```

```

<400> 78
ggaggttcag gaagccattt tctttcgcgt ctgcggtgct cggagtgtgg tacttctcct      60
agttgcagtc aggttcata cgctattgtc ctgcccgta gagcagccag cgggtacaga      120
atggattttg gaagagggag tcaccactgg acctccaagg aagccacgtg cagacatcta      180
caaccttcga tctcctgacg agtttattgt tggccaaaac caggctttga ttgaaccagg      240
atgaatgcgg gtgttggaag tagaatatat atatacatat aaaattgggtt gggagccacg      300
tgtaccagtg tgtgttgatc ttggcttgat tcagtctgcc ttgtaacaga aactggcgat      360
ggaatatgag aggagccctc tggaaagaaa aggacagacc ctgtgctttc atgaaagtga      420
agatctggct gaaccagttc cacaagggtta ctgtatacat agcctgagtt taaaaggctg      480
tgcccacttc aagaatgtca ttgttagact ttgaaatttc taactgccta cctgcataaa      540
gaaaataaaa tcttttaaat caaaaaaaaa aacaaaaaaa aatatctata aacacaaaga      600
aaaacaaatt caaacttaca cattatccaa tcattctatt attataaata tatgaaaaaa      660

```

121

aaaaaaactt gggggcccg c ttggcccagg gggtttttaa tctggaccaa tttattcttg 720
 gggcccaatt tggggctcat tttggccaat tgccccttta ggggaggcgt ttacaattca 780
 tgggcccgtt tttcaaacgg tggtatttgg aaaaccctg gggttacca atttttaaac 840
 ccttgggaaa aaatnccctt ttttcacagt ggggggaaat tccgaaggcc ccgcccctg 900
 aggccttttc catttggecc ctctttgg 928

<210> 79
 <211> 2448
 <212> DNA
 <213> Homo sapien

<400> 79
 tttttttttt tcactagtag cattgttaat attaaacaag atagctaata gcttttaaagc 60
 gtctaatacac ggtaccccca cgagtagtag tcacacagca aatcataggc aggagtggac 120
 aaacagcgct gggagcatta attctggtag atttgagtc tgggcagggc ctaagaggat 180
 gggatcagcg gattgggcca gccaaagggg gcggttcag agggaggagc cagctgcctc 240
 ccgctgcaag ccccgagctt tgaatcaagc gcggtagttt cagctgtagc caatcctaga 300
 acaaaatctt tttgcgtcag ttgttgctga ggcctagatc tcaccagat cagttctagt 360
 cgctgctgct tgtaccgctc ctggccctt gcgtgtccgt ccgttacgac tcagaggacg 420
 gagagaaagt ctaagtcacc ttctcagta tccaggccgg aaggagactc tcagctactc 480
 gccctgaaca tgcagagcag cactggcttc cctgtctctg gatgccgcag gggcgtcctg 540
 tcaggagttg tagtgtcaag attccagttc ttcttatgcg acttgctgga gaagcaggaa 600
 gttgctgttt attggagaag tgggctagg gagggaaaca atcccgcgga gccatatgga 660
 gagggcagga ggtgtctaga tagtggtagg ccggcggggg acggagggta tgggtggtga 720
 gtcctggccc tttaaacccg agggggagga tccggaagtg gatgggcggg gcgaagggtc 780
 ctggtatata aaaggccccc ggccgcgcgg ggtctctaatt ctgccatctt ctgtccctga 840
 gtgagtctct ggcgtcccaa attgcctgtt tttctcgag gctctattcc gttecgctgg 900
 tgcacacctc aggggaacga tggccatgga gtccacagcc actgccgccg tgcgcgcgga 960
 gctggtttct gccgacaaaa ttgaagatgt tctgtctct tctacatctg cagataaagt 1020
 ggagagcaac cattaatctg ctttctgtca ctgaagatgg gttgcatttt gtagaatatt 1080
 atctaaatag aatcatacat ctggatgtgg atagtgaagc taagaaacta ttgggttttag 1140
 gacagaaaca tctggtgatg ggggatattc cagcagctgt caatgcattc caggaagcag 1200
 ctagtctttt aggtaagaag tatggagaga cagctaata gtgtggagaa gccttctttt 1260
 tctatgggaa atcacttctg gagttggcaa gaatggagaa tgggtgtgtg ggaaacgcct 1320

122

tggaaggtgt gcatgtggaa gaggaagaag gagaaaaaac agaagatgaa tctctggtag 1380
 aaaataatga taacatagat gaaactgaag gctcggaaga ggatgataaa gaaaatgata 1440
 agactgaaga aatgccaaat gattcagtcc ttgaaaacaa gtctcttcaa gaaaatgagg 1500
 aggaggagat tgggaaccta gagcttgccct gggatatgct ggatttagca aagatcattt 1560
 ttaaaaggca agaaacaaaa gaagcacagc tttatgctgc ccaggcacat cttaaactcg 1620
 gagaagttag tgttgaatct gaaaactatg tgcaagctgt ggaggagtgc cagtcctgcc 1680
 ttaacctgca ggaacagtac ctggaagccc acgaccgtct ccttgcagag acccactacc 1740
 agctgggctt ggcttatggg tacaactctc agtatgatga ggcagtggca cagttcagca 1800
 aatctattga agtcattgag aacagaatgg ctgtactaaa cgagcaggtg aaggaggctg 1860
 aaggatcgtc tgctgaatac aagaaagaaa ttgaggaact aaaggaactg ctacccgaaa 1920
 ttagagagaa gatagaagat gcaaaggagt ctacgcgtag tgggaatgta gctgaactgg 1980
 ctctgaaagc tactctggtg gagagttcta cttcagggtt cactcctggt ggaggaggct 2040
 cttcagtctc catgattgcc agtagaaagc caacagacgg tgcttcctca tcaaattgtg 2100
 tgactgatat ttcccacctt gtcagaaaga agaggaaacc agaggaagag agtccccgga 2160
 aagatgatgc aaagaaagcc aaacaagagc cggaggtgaa cggaggcagt ggggatgctg 2220
 tccccagtgg aaatgaagtt tcggaaaaca tggaggagga ggctgagaat caggctgaaa 2280
 gccgggcagc agtggagggg acagtggagg ctggagctac agttgaaagc actgcatgtt 2340
 aagagggggc acagccctcc tccaaggga aagtgttttt gtatataatg tattttttca 2400
 cttttggagg attctttttg tataacttca ataaagattg taagcaaa 2448

<210> 80
 <211> 1648
 <212> DNA
 <213> Homo sapien

<400> 80
 tgacgcgaat gacgctgcgc cagtcaggcc gcagccccgc tgcttgcccc gtcgggcccc 60
 ccttgccccg ctcgccccgc ccccgccct cctgcacgg cctcccgctgc gcccctgtca 120
 gactgtggcg gccggtcgcg cgggtgcgctc tccctccctg cccgcagcct ggagaggcgc 180
 ttctgtctgc acacccccgc gttcctgccc gcaccgcgcc tgccctctgc cgcgctccgc 240
 cctgcccgcg accgcacgcc cgcgcgggga catggcacac gcaccggcac gctgccccag 300
 cgcgccggggc tccggggacg gcgagatggg caagcccagg aacgtggcgc tcatcaccgg 360
 tatcacaggc caggatggtt cctacctggc tgagttcctg ctggagaaag gctatgaggt 420
 ccatggaatt gtacggcggc ccagttcatt taatacgggt cgaattgagc atctgtataa 480

123

gaatccccag gctcacattg aaggaaacat gaagttgcac tatggcgatc tcaactgacag 540
 tacctgcctt gtgaagatca ttaatgaagt aaagcccaca gagatctaca accttgaggc 600
 ccagagccac gtcaaaattt cctttgacct cgctgagtac actgcggacg ttgacggagt 660
 tggcactcta cgacttctag atgcagttaa gacttgtggc cttatcaact ctgtgaagtt 720
 ctaccaagcc tcaacaagtg aactttatgg gaaagtgcag gaaatacccc agaaggagac 780
 cacccttttc tatccccggg caccctatgg agctaatttc gttactcgaa aaattagccg 840
 gtcagtagct aagattttacc ttggacaact ggaatgtttc agtttgggaa atctggatgc 900
 caaacgagat tggggccatg ccaaggacta tgtggaggct atgtgggtga tgttgcagaa 960
 tgatgagccg gaggacttcg ttatagctac tggggaggct catagtgtcc gggaatttgt 1020
 cgagaaatca ttcttgcaca ttggaaaaac cattgtgtgg gaaggaaaga atgaaaatga 1080
 agtgggcaga tgtaaagaga ccggcaaagt tcacgtgact gtggatctca agtactaccg 1140
 gccaaactgaa gtggactttc tgcagggcga ctgcaccaa gcgaaacaga agctgaactg 1200
 gaagccccgg gtcgctttcg atgagctggg gagggagatg gtgcacgccg acgtggagct 1260
 catgaggaca aacccaatg cctgagcagc gcctcggagc ccggcccgcc ctccggctac 1320
 aatccccgca gagtctccgg tgcagacggc ctgcggggat ggggagcggc gtgccaatct 1380
 gcgggtcccc tgcggccctt gctgccgctg cgctgtcccg gccgcaagag cggggccgccc 1440
 ccgcccagggt ttgtagcagc cgggatgtga cctccagggt tttgggtcgc tttgcgtttg 1500
 tcgaagcctc ctctgaatgg ctttgtgaaa tcaagatgtt ttaatcacat tcactttact 1560
 tgaaattatg ttgttacaca acaaattgtg gggccttcaa attgtttttc tcttttcata 1620
 ttaaaaatgg tctttctgtg aactagca 1648

<210> 81
 <211> 2595
 <212> DNA
 <213> Homo sapien

<400> 81
 tgacgcgaat gacgtgcgc cagtcaggcc gcagccccgc tgcttggccc gtcgggcccc 60
 ccttggccgg ctcggccccgc ccccgccctt ccctgcacgg cctcccggtc gccctgtca 120
 gactgtggcg gccggtcgcg oggtgcgctc tccctccctg cccgcagcct ggagaggcgc 180
 ttcgtgctgc acacccccgc gttcctgccc gcaccgcgcc tgccctctgc cgcgctccgc 240
 cctgccgccc accgcacgcc cgcgcgggga catggcacac gcaccggcac gctgccccag 300
 cgccccgggg tccggggacg gcgagatggg caagcccagg aacgtggcgc tcatcaccgg 360
 tatcacaggg caggatgggt cctacctggc tgagttcctg ctggagaaaag gctatgaggt 420

ccatggaatt gtacggcggg ccagttcatt taatacgggt cgaattgagc atctgtataa	480
gaatccccag gctcacattg aaggaaacat gaagttgcac tatggcgatc tcaactgacag	540
tacctgcctt gtgaagatca ttaatgaagt aaagcccaca gagatctaca acctgggagc	600
ccagagccac gtcaaaattt cctttgacct cgctgagtac actgcggacg ttgacggagt	660
tggcactcta cgacttctag atgcagttaa gacttgtggc cttatcaact ctgtgaagtt	720
ctaccaagcc tcaacaagtg aacttttatgg gaaagtgcag gaaatacccc agaaggagac	780
cacccttttc tatccccggg caccctatgg ggcagcaaaa ctctatgcct attggattgt	840
ggatgaacttc cgtgaggcgt ataatctctt tgcagtgaac ggcattctct tcaatcatga	900
gagtcccaga agaggagcta atttcgttac tcgaaaaatt agccggtcag tagctaagat	960
ttaccttgga caactggaat gtttcagttt gggaaatctg gatgccaaac gagattgggg	1020
ccatgccaaag gactatgtgg aggctatgtg gttgatgttg cagaatgatg agccggagga	1080
cttcgttata gctactgggg aggtccatag tgtccgggaa tttgtcgaga aatcattctt	1140
gcacattgga aaaaccattg tgtgggaagg aaagaatgaa aatgaagtgg gcagatgtaa	1200
agagaccggc aaagttcacg tgactgtgga tctcaagtac taccggccaa ctgaagtgga	1260
gaccaattaa ttgttcaaag gtaatatata gagatctcaa gagcatttac tttggcacag	1320
tctctgcccc tttgtggtct ggctaacaat aaagctaaga tttgttttgt aatgacggtt	1380
ccagttggac tggaattcct ttgaagttcc tgaagtttgc tgttgtgaga gctcaagata	1440
ttatttgact gcaatgaaac actctctgtg tattatcaaa atatcgtata cactgaactg	1500
taacaacttg cagtaatcag acttctcctt cgagggtggc tgcgtgagta gagagcatcc	1560
acagagagat gagcttctcc caacatccat ttagtgttca gtcacacgg gccactcatg	1620
gatctgaacc ttcttacaag gtcgatacca acaggtatga ggagcagaga gcagatcagc	1680
acggagcagc attagctggg gcctccttgc ctggcctggc agtgaagata gagcaccatg	1740
ccagttgttc agcttctcag cgcttgcatc agctgacct aattatttaa ggatgttgtc	1800
acctctaaat aaatatagtc attgtcctgg tctctattta tatttttaaaa ctctttccca	1860
gcccaccccg atggcctaaa acaattcata tacttcagta gaagagggtg tcacagggt	1920
atttgcacag gtttggatcc agtgccagcg ttagccttca cacaagttac actcctctga	1980
gatgtaactg tctcctgtgt atgagtgcc aagtatttga taaccgccat ctgtctgctc	2040
tcaacacagg tttcacaagc acagtgatgt cttccattgt gtgcgtaaca actgtgtgtg	2100
tgcggagaga gaggactctg tacgcatgca catgtttgga gctacataaa tctgagaagc	2160
agttttgtct agtgtcgttt cagtaaattgg gactcattta gtcagtttgc tttgctgtca	2220

125

```

gaatcatatt ccagatatca gaatgtgcta tggcagggtac ttgtttaaaa caagtaataa 2280
tatccttgta atgttttagca atacagctga ctgggtctgtg ttccattact tgtgtaaaag 2340
gaaatatgta aagcttttaa aaataagttt aaatgatcaa aagagatgcc tgaaaatata 2400
ttttggcaca ggaaattaca ggaatttata aatgtgtcat gtggaaagat aattacaaag 2460
tttatttttaa tcttttatct ctgactgaat cttaagttaa aatattcttc ccggcatcct 2520
gtctggtttt tgccactgtc cctgctttac gtgttacctt tccatgttga tctgagttgc 2580
atttagttga ggaga 2595

```

<210> 82

<211> 1706

<212> DNA

<213> Homo sapien

<400> 82

```

tgacgcgaat gacgctgcgc cagtcaggcc gcagccccgc tgcttgcccc gtcgggcccc 60
ccttgccccg ctcggccccg ccccgccctt ccctgcacgg cctcccgtgc gcccctgtca 120
gactgtggcg gccggtcgcg cgggtgcgtc tccctccctg cccgcagcct ggagaggcgc 180
ttcgtgctgc acacccccgc gttcctgccg gcaccgcgcc tgccctctgc cgcgctccgc 240
cctgccgccg accgcacgcc cgcgcgggga catggcacac gcaccggcac gctgccccag 300
cgccccgggg tccggggacg gcgagatggg caagcccagg aacgtggcgc tcatcaccgg 360
tatcacaggc caggatgggt cctacctggc tgagttcctg ctggagaaag gctatgaggt 420
ccatggaatt gtacggcggg ccagttcatt taatacgggt cgaattgagc atctgtataa 480
gaatccccag gtcacattg aaggaaacat gaagttgcac tatggcgatc tcaactgacag 540
tacctgcctt gtgaagatca ttaatgaagt aaagcccaca gagatctaca accttgagac 600
ccagagccac gtcaaaattt cctttgacct cgctgagtac actgcggacg ttgacggagt 660
tggcactcta cgacttctag atgcagttaa gacttgtggc cttatcaact ctgtgaagtt 720
ctaccaagcc tcaacaagtg aactttatgg gaaagtgcag gaaatacccc agaaggagac 780
caccctttc tatccccggt caccctatgg ggcagcaaaa ctctatgcct attggattgt 840
gggtgaacttc cgtgaggcgt ataatctctt tgcagtgaac ggcattctct tcaatcatga 900
gagtccaga agaggagcta atttcgttac tcgaaaaatt agccggtcag tagctaagat 960
ttaccttgga caactggaat gtttcagttt gggaaatctg gatgccaaac gagattgggg 1020
ccatgccaaag gactatgtgg aggctatgtg gttgatgttg cagaatgatg agccggagga 1080
cttcgttata gctactgggg aggtccatag tgtccgggaa tttgtcgaga aatcattctt 1140
gcacattgga aaaaccattg tgtgggaagg aaagaatgaa aatgaagtgg gcagatgtaa 1200

```

126

agagaccggc aaagttcacg tgactgtgga tctcaagtac tacgggcca ctgaagtggc 1260
 aaggactctc tggccacca gtgcgtggcc acgtttggct ggctggctgg gcaaattgcgc 1320
 tcatggaatg cctgggtgcca gcctttggag ctgtcagttc agctccctgg catccttcag 1380
 cgtgcatttc caaaatcaga acacagtga ttcaatttga gattatcatt cttaccacaa 1440
 acttcattgc tctataaagg aaaacaagag ttaaaccac caaaagaaat cattgatttc 1500
 agaaatggaa tctcacatca cacgcagatc tctcatgtt ggtctcctat tctaattggc 1560
 ggtccctgat aacctacagg atgcctgact taacaacgct gtcccagacc tgcaggacat 1620
 ctaggtctc tccagctatg tctatgtctg tgtctgtgta ctcatacaca gtgcatactt 1680
 gcacactata ccataaatgg tctccc 1706

<210> 83
 <211> 2667
 <212> DNA
 <213> Homo sapien

<400> 83
 ttggtttttt tttttttttt gcttttttgt tttttgagac agggctcttg tctgttgccc 60
 aggtctgggg gtagtggcga gatctcagtt cactgcaacc tccacatccc ggctttaacc 120
 gattctccca tctcagcctc acgagtagct gggattacag ggctgcatca acacacctgg 180
 ctaatttttg tatttttagt acagatgggg tttcaccatg ttggccaggc tggctctcaa 240
 ctctgacct caggtgatct gccagcttca gcctcccaa gtgctgggat tacaggcgtg 300
 agctaccacg cctggccttg gccatgtaga tttctttcac accttcctta ttccttcttc 360
 agaagctgag gcccacatgag ggaagctgg tagccagggg tcttttcagg tacctgtggc 420
 caaattaggc ccagatgtta atgctctcat ccatggctga aaaaggccaa gaaactaaga 480
 agagctgtgg acgcacagac tggcagagac accgagaggg cactgcgacc gcactgcaca 540
 gatgagcaaa ggaaggccaa cttcagcatt ccctcctcct ctctaagccc ctccaaaagg 600
 cctgctcagg gagctctggc caccgttctc aagcaattgc ctgacgtttt tactgacact 660
 gagggctgga gatggtccct tttccaggca gctatgtgtc catttcctag agagggaaac 720
 tgagtccag agaaggcggg tgtggctgga ggccatgggt cagtcatagg cataaccatg 780
 ccgtgaataa ccgtcttggc actctctgac ttctttccct ggcaatacaa cggatcccag 840
 tgcccatttg tcacataagg aaactgaggg tggcgtcttc gatgaagcct ggcaggagtg 900
 ggggtcacac acggagtgca aggtgtcca agccttcaag gtcattctga ctgatctttg 960
 agccttagca atacaggacg tgggtccaag agcatttcca gcacaaaatg gcgcctcata 1020
 gaaaaatgta gtcggagttc cagcgccac ctcaccacg gctcaaagga gggggcgcgg 1080

127

gcggcagcgc acacttcatt tgcataataca aaaatcaacg gatttcggtg gctggcagcg 1140
 gtttgatctg cacgaaactg ggctcgccctc tgaccggtct ctatcagtca ttggcccatg 1200
 cgagtctcgt gacccaccag tcaactaaagc gcatccctat tggcccgagc gccttccatc 1260
 acgcgtttta aacccccgagt cccgcctccc ctcgccacgc tccggccccc aggaagaccg 1320
 ttagtcgcgt gcgcgacttc ctcagcgtca gggcgcgccc gcgcgtatgc cgggcccgtg 1380
 cggcgtctca ttcaaacggt ccaatcagcg gccgcgtcgc cttccccgtc ccccttctc 1440
 tcttcttccc cttttcagcc cctccgttct ttgagaacct tcccgacccc cgcgggcccc 1500
 cattggccgg cccgctaccc agccctctcc gccacttccc tcgcttctga ccatagtttg 1560
 cggggaaggg agcgagcgcg tcgaaaacca aggaacgtgc gcgctgacgt cacggttgag 1620
 gctcggagct gaggggcccgc ggagggcgtg gcctgcgggc ggttataaag aggcagtgg 1680
 gcgcgcgcgg ccggctcagt gctgccgggc accggggcgg cgggttggtc tacgctgtgc 1740
 gcggcggacg tcggaggcag cggggagcgg agcggggccg cgggggcctc tccagggccg 1800
 cagcggcagc agttgggccc cccgccccgg ccggcggacc gaagaacgca ggaagggggc 1860
 cggggggacc cgccccggc cggccgcagc catgaactcc aacgtggaga acctacccc 1920
 gcacatcatc cgcctggtgt acaaggaggt gacgacactg accgcagacc caccgatgg 1980
 catcaaggtc tttcccaacg aggaggacct caccgacctc caggtcacca tcgagggccc 2040
 tgaggggacc ccatatgctg gaggtctgtt ccgcatgaaa ctctgctgg ggaaggactt 2100
 ccctgcctcc ccaccaagg gctacttctc gaccaagatc ttccacccga acgtgggccc 2160
 caatggcgag atctgcgtca acgtgctcaa gagggactgg acggtgagc tgggcatccg 2220
 acacgtactg ctgaccatca agtgctgct gatccaccct aaccccgagt ctgcactcaa 2280
 cgaggaggcg ggccgcctgc tcttgagaaa ctacgaggag tatgcagctc gggcccgtct 2340
 gctcacagag atccacgggg gcgcggcgcc gccagcggc agggccgaag ccggtcgggc 2400
 cctggccagt ggcaactgaag ctctctccac cgaccctggg gcccggggcc ggagggctga 2460
 agtccattgg ccagagcatg tgggaaggga aagacgttgg ggcagaaaaga cgacagacgg 2520
 cgcgcgcgtg aaggtctttc tctcacgtga ccattctgct cccaatttta gtaattgcgg 2580
 cccgagcggg cgccgtgtca atgccccaac aaaaaaacct tctcgaaaag gtgtcctcag 2640
 tgctgctttt caggcgtcgc ttcttca 2667

<210> 84

<211> 2648

<212> DNA

<213> Homo sapien

<400> 84

128

ttgggtttttt tttttttttt gcttttttgt tttttgagac aggggtcttgc tctgttgccc	60
aggctgggggt gcagtggcga gatctcagtt cactgcaacc tccacatccc ggctttaacc	120
gattctccca tctcagcctc acgagtagct gggattacag ggctgcatca acacacctgg	180
ctaatttttg tatttttagt acagatgggg tttcaccatg ttggccaggc tgggtctcaa	240
ctcctgacct caggatgatct gccagcttca gcctcccaa gtgctgggat tacaggcgtg	300
agctaccacg cctggccttg gccatgtaga tttctttcac accttcctta ttcttcttc	360
agaagctgag gccccatgag ggaagctgg tagccagggg tcttttcagg tacctgtggc	420
caaattaggc ccagatgtta atgctctcat ccattggctga aaaaggccaa gaaactaaga	480
agagctgtgg acgcacagac tggcagagac accgagaggg cactgcgacc gcactgcaca	540
gatgagcaaa ggaaggccaa cttcagcatt cctcctcct ctctaagccc ctccaaaagg	600
cctgctcagg gagctctggc caccgttctc aagcaattgc ctgacgtttt tactgacact	660
gagggctgga gatggctcct tttccaggca gctatgtgtc catttcctag agagggaaac	720
tgagtcccag agaaggcggg tgtggctgga ggccatgggt cagtcatagg cataaccatg	780
ccgtgaataa ccgtcttggc actctctgac ttctttccct ggcaatacaa cggatcccag	840
tgccatttg tcacataagg aaactgaggc tggcgtcttc gatgaagcct ggcaggagtg	900
ggggtcacac acggagtgca aggctgtcca agccttcaag gtcattctga ctgatctttg	960
agccttagca atacaggacg tgggtccaag agcatttcca gcacaaaatg ggcctcata	1020
gaaaaatgta gtcggagtgc cagcgccac ctccaccacg gctcaaagga gggggcgcg	1080
gcggcagcgc acacttcatt tgcataatac aaaatcaacg gatttcgttg gctggcagcg	1140
gtttgatctg caggaactg ggtcgccctc tgaccggtct ctatcagtca ttggcccatg	1200
cgagtctcgt gacccaccag tcaactaaagc gcattccctat tggcccgagc gccttcctc	1260
acgcgtttaa aaccccgagt cccgcctccc ctcgccacgc tccggcccc aggaagaccg	1320
ttagtgcgt gcgcgacttc ctcagcgtca gggcgcgccc gcgcgtatgc cgggccgtg	1380
cggcgtctca ttcaaacggt ccaatcagcg gccgcgtcgc cttccccgtc ccccttcctc	1440
tcttcttccc cttttcagcc cctccgttct ttgagaaccc tcccgacccc cgcgggcccc	1500
cattggccgg cccgctaccc agccctctcc gccacttccc tcgcttctga ccatagtttg	1560
cggggaagg agcgagcgc tcgaaaacca aggaacgtgc gcgctgacgt cacggttgag	1620
gctcggagct gaggggccc gcggggcggt gcctgcgggc ggttataaag aggcagtgg	1680
gcgcgcgcgc ccggctcagt gctgccgggc accggggcg cgggttggtc tacgctgtgc	1740
gcggcggacg tcggaggcag cggggagcgc agcggggcg ccggggcctc tccagggccg	1800
cagcggcagc agttgggccc cccgccccgc ccggcggacc gaagaaccta ccccgacaca	1860

129

```

tcatccgcct ggtgtacaag gaggtgacga cactgaccgc agaccacccc gatggcatca 1920
aggtctttcc caacgaggag gacctaccg acctccaggt caccatcgag ggcctgagg 1980
ggaccccata tgctggagggt ctgttccgca tgaaactcct gctggggaag gacttccctg 2040
cctccccacc caagggctac ttcctgacca agatcttcca cccgaacgtg ggcgccaatg 2100
gcgagatctg cgtcaacgtg ctcaagaggg actggacggc tgagctgggc atccgacacg 2160
tactgctgac catcaagtgc ctgctgatcc accctaacc cgagtctgca ctcaacgagg 2220
aggcgggccc cctgctcttg gagaactacg aggagtatgc agctcgggccc cgtctgctca 2280
cagagatcca cggggcgccc ggcgggcccc gcggcagggc cgaagccggg cgggcccctgg 2340
ccagtggcac tgaagcttcc tccaccgacc ctggggcccc agggggcccc ggaggggctg 2400
aggggtcccat ggccaagaag catgctggcg agcgcgataa gaagctggcg gccaaagaaa 2460
agacggacaa gaagcgggcg ctgcggcggc tgtagtgggc tctcttcctc cttccaccgt 2520
gaccccaacc tctcctgtcc cctccctcca actctgtctc taagttatct aaattatggc 2580
tggggtcggg gagggtagag ggggcactgg gacctggatt tgtttttcta aataaagttg 2640
gaaaagca 2648

```

```

<210> 85
<211> 2800
<212> DNA
<213> Homo sapien

```

```

<400> 85
ttggtttttt tttttttttt gcttttttgt tttttgagac aggggtcttg tctgttgccc 60
aggtctggggg gcagtggcga gatctcagtt cactgcaacc tccacatccc ggctttaacc 120
gattctccca tctcagcctc acgagtagct gggattacag ggctgcatca acacacctgg 180
ctaatttttg tatttttagt acagatgggg ttccaccatg ttggccaggc tggctctcaa 240
ctcctgacct caggatgatc gccagcttca gcctcccaaa gtgctgggat tacaggcgtg 300
agctaccacg cctggccttg gccatgtaga tttctttcac accttcctta ttccttcttc 360
agaagctgag gccccatgag ggggaagctg tagccagggg tcttttcagg tacctgtggc 420
caaattaggc ccagatgtta atgctctcat ccatggctga aaaaggccaa gaaactaaga 480
agagctgtgg acgcacagac tggcagagac accgagaggg cactgcgacc gcactgcaca 540
gatgagcaaa ggaaggccaa cttcagcatt cctcctcct ctctaagccc ctccaaaagg 600
cctgctcagg gagctctggc caccgttctc aagcaattgc ctgacgtttt tactgacact 660
gagggctgga gatggctcct tttccaggca gctatgtgtc catttcctag agaggggaaac 720
tgagtcccag agaaggcggg tgtggctgga ggccatgggt cagtcatagg cataaccatg 780

```

ccgtgaataa ccgtcttggc actctctgac ttctttccct ggcaatacaa cggatcccag 840
 tgcccatttg tcacataagg aaactgaggc tggcgtcttc gatgaagcct ggcaggagtg 900
 ggggtcacac acggagtgca aggctgtcca agccttcaag gtcattctga ctgatctttg 960
 agccttagca atacaggacg tgggtccaag agcatttcca gcacaaaatg gcgcctcata 1020
 gaaaaatgta gtcggagttc cagcgcccac ctccaccacg gctcaaagga gggggcgcg 1080
 gcggcagcgc acacttcatt tgcatataca aaaatcaacg gatttcggtg gctggcagcg 1140
 gtttgatctg cacgaaactg ggctgccttc tgaccggtct ctatcagtca ttggcccatg 1200
 cgagtctcgt gaccaccacg tactaaagc gcatccctat tggcccgagc gccttccatc 1260
 acgcgtttaa aaccccgagt cccgcctccc ctcgccacgc tccggccccc aggaagaccg 1320
 ttagtcgctg gcgcgacttc ctacagctca gggcgcgccc gcgcgtatgc cgggcgctgg 1380
 cggcgctctca ttcaaacggt ccaatcagcg gccgcgtcgc cttccccgtc ccccttccctc 1440
 tcttcttccc cttttcagcc cctccgttct ttgagaacct tcccgacccc cgcgggcccc 1500
 cattggccgg cccgctaccc agccctctcc gccacttccc tcgcttctga ccatagtttg 1560
 cggggaaggg agcgagcgcg tcgaaaacca aggaacgtgc gcgctgacgt cacggttgag 1620
 gctcggagct gaggggcccgc ggagggcgtg gcctgcgggc gggtataaag aggcagtgg 1680
 gcgcgcgcgg ccggctcagt gctgccgggc accggggcgg cgggttggtc tacgctgtgc 1740
 gcggcggaag tcggaggcag cggggagcgg agcggggccg ccggggcctc tccagggccg 1800
 cagcggcagc agttgggccc cccgccccgg ccggcggaac gaagaacgca ggaagggggc 1860
 cggggggacc cggccccggc cggccgcagc catgaactcc aacgtggaga acctaccccc 1920
 gcacatcatc cgcttgggtg acaaggaggt gacgacactg accgcagacc caccgatgg 1980
 catcaaggtc tttcccaacg aggaggacct caccgacctc caggtaacca tcgagggcc 2040
 tgaggggacc ccatatgctg gaggtctgtt ccgcatgaaa ctctgctgg ggaaggactt 2100
 ccctgcctcc ccaccaagg gctacttctt gaccaagatc ttccacccga acgtgggcgc 2160
 caatggcgag atctgcgtca acgtgctcaa gagggactgg acggctgagc tgggcatccg 2220
 acacgtactg ctgctttcct ggaaagacaa acagtgcag acccaggaca cccaagtact 2280
 gttgagaagt gccaggagc accttactat gcaacgagt accatcaagt gcctgctgat 2340
 ccaccctaac cccgagtctg cactcaacga ggaggcgggc cgctgctct tgagaaacta 2400
 cgaggagtat gcagctcggg cccgtctgct cacagagatc cacgggggcg ccggcggggc 2460
 cagcggcagg gccgaagccg gtcgggccct ggccagtggc actgaagctt cctccaccga 2520
 ccctggggcc ccagggggcc cgggaggggc tgagggtccc atggccaaga agcatgctgg 2580

131

cgagcgcgat aagaagctgg cggccaagaa aaagacggac aagaagcggg cgctgcggcg 2640
 gctgtagtgg gctctcttcc tccttccacc gtgaccccaa cctctcctgt cccctccctc 2700
 caactctgtc tctaagttat ttaaattatg gctggggtcg gggaggggtac agggggcact 2760
 gggacctgga ttgttttttc taaataaagt tggaaaagca 2800

<210> 86
 <211> 717
 <212> DNA
 <213> Homo sapien

<400> 86
 cccaggggc tcatcccagc ctggacttga cacgttgctc tctgttttta cctaaaagag 60
 ccagaagcgc tatcacgaag atatctttgg tgctgtattt ccttacgagg cgaagaagga 120
 caggggtggcg gtgcagccca gcagcctgga gatgtcagcg ctctgagggc ctacaggagg 180
 ggttaaagct gccggcacag aacttaagga tggagccagc tctgcattat ctgaggtcac 240
 agggcctggg gagatggagg aaagtgatat cccccagcct caagtcttat ttctcaacg 300
 ttgtcccca tcaagccctt tacttgacct cctaacaagt agcaccctgg attgatcgga 360
 gcctcctctc tcaaaactggg gcctccctgg tcccttgagg acaaaatctt aaatgccagg 420
 cctggcaagt gggtgaaaga tggaaacttg tgctgagtgc accacttcaa gtgaccacca 480
 ggaggtgcta tcgcaccact gtgtatttaa ctgccttggtg tacagttatt tatgcctctg 540
 tatttaaaaa actaacaccc agtctgttcc ccatggccac ttgggtcttc cctgtatgat 600
 tccttgatgg agatatttac atgaattgca ttttacttta atcacactgt atgcgtgtgt 660
 ggggtgttttg tagggaaagc tcttctcaga gtggggagct ggtgggtgtc acagcct 717

<210> 87
 <211> 1098
 <212> DNA
 <213> Homo sapien

<400> 87
 tgccagggag ctccaacaga actgaggagc gaacagtagg tgccctgagag ctggccgatc 60
 tggtcactct ctccagtggg ccacaggggc cacttggtg ggcagtctct gctgccagct 120
 agggcatctc agctcagccc agagcaaaga cgtgtagggg cataacgctt tcctcaagct 180
 tcctttccca ggtccaaact tagaggccca gttttgggccc tagaattcac tgaggtttgt 240
 ggctgtttgt ccaacagtca ccatggatca ccgaagccga ctacggggca caggcctgaa 300
 ccgaatccct gggactcagt cccgagcccc ccgagtccca ctccccttcc acgtgcaaca 360
 ggaggccagg gaaggagaag actgggagcg agagccacct cgtcagaggc ctcctatcta 420
 tgagccacca gaaagtgaag agctgccaga taatgttatg gtttccaagc cagcgccta 480

132

```

ctgggaagga acagctgtga tcgatggaga atttaaggag ctgaagttaa ctgattatcg      540
tgggaaatac ttggttttct tcttctaccc acttgatttc acatttgtgt gtccaactga      600
aattatcgct tttggcgaca gacttgaaga attcagatct ataaatactg aagtggtagc      660
atgctctgtt gattcacagt ttacccattht ggctggatt aatacccctc gaagacaagg      720
aggacttggg ccaataagga ttccacttct ttcagatttg acccatcaga tctcaaagga      780
ctatggtgta tacctagagg actcaggcca cactcttaga ggtctcttca ttattgatga      840
caaaggaatc ctaagacaaa ttactctgaa tgatcttctt gtgggtagat cagtggatga      900
gacactacgt ttggttcaag cattccagta cactgacaaa cacggagaag tctgccctgc      960
tggtctggaaa cctggtagtg aaacaataat cccagatcca gctggaaagc tgaagtattt    1020
cgataaactg aattgagaaa tacttcttca agttatgatg cttgaaagtt ctcaataaag    1080
ttcacggttt cattacca                                          1098

```

```

<210> 88
<211> 731
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (21)..(21)
<223> n=a, c, g or t

```

```

<400> 88
tatagataga tatgtttatc nccataactc ataactggga gtattatacc caagaggctt      60
tgttacaagg tatttttttac cctgttagta aatattttcc tgtaattcta taataagcaa    120
ataaattttc atggttgggg tagatgcaaa aaaaaagaa ttcgaattcg aattcggaac    180
gtcatcgttt ggaaagcgtc gcaataagac gcacacgttg tgccgccgct gtggctctaa    240
ggcctaccac cttcagaagt cgacctgttg caaatgtggc taccctgcca agcgcaagag    300
aaagtataac tggagtgcc aaggctaaaag acgaaatacc accggaactg gtcgaatgag    360
gcacctaaaa attgtatacc gcagattcag gcatggattc cgtgaaggaa caacacctaa    420
accaagagg gcagctgttg cagcatccag ttcattctaa gaatgtcaac gattagtcac    480
gcaataaatg ttctggtttt aaaaaataca tatctggttt tggttaaggta tttttaatca    540
attaggcttg tagtatcagt gaaatactgt aggtttaggg actgggctag cttcatatca    600
gatttacttg ttaagtgact gttttggaat gtttactttt ggactgggtt tgtaacacgg    660
ttaaaggcaa tgagaaacaa gcagaattcc aggagtcctt gaagcagagg gcactggaag    720
acaatatagc a                                          731

```

133

<210> 89
<211> 777
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (50)..(50)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (90)..(90)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (199)..(199)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (201)..(201)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (204)..(204)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (206)..(206)
<223> n=a, c, g or t

<400> 89
atgttctttg ttgacaatc tcataagact ttagtcttac agacgaccn tcgaagtctg 60
ctggcacaca ccacatgcag attttggtgn tttcccaacc ttcttggtat aaaggtaaac 120
aattctatta ccaggggttc gggacagcct agttttgtta gaggctgtat tgtaggaaag 180
cctacgtcgg tatgtcaana nctntncgaa ttcggecgag gggaacgtca tcgtttggaa 240
agcgtcgcaa taagacgcac acgttggtgcc gccgctgtgg ctctaaggcc taccaccttc 300
agaagtcgac ctgtggcaaa tgtggctacc ctgccaaagcg caagagaaag tataactgga 360
gtgccaaggc taaaagacga aataccaccg gaactggtcg aatgaggcac ctaaaaattg 420
tataccgcag attcagcat ggattccgtg aaggaacaac acctaaacc aagagggcag 480
ctgttcgagc atccagttca tcttaagaat gtcaacgatt agtcatgcaa taaatgttct 540

134

gggttttaaaa aatacatatc tgggttttggg aaggtatttt taatcaatta ggcttgtagt 600
 atcagtga aa tactgtagg ttagggactg ggctagcttc atatcagatt tacttgtaaa 660
 gtgactgttt tggaatgttt acttttggac tgggtttgta acacgggttaa aggcaatgag 720
 aaacaagcag aattccagga gtccttgaag cagagggcac tgggaagaaa tatagca 777

<210> 90
 <211> 564
 <212> DNA
 <213> Homo sapien

<400> 90
 atgcatgccg agcggcgagc tgtgatggat cgtggtcgag gccgaggtcg ttgtgcccgc 60
 gctgtggctc taaggcctac caccttcaga agtcgacctg tggcaaatgt ggctaccctg 120
 ccaagcgcaa gagaaagtat aactggagtg ccaaggctaa aagacgaaat accaccggaa 180
 ctggtcgaat gaggcaccta aaaattgtat accgcagatt caggcatgga ttccgtgaag 240
 gaacaacacc taaacccaag agggcagctg ttgcagcctc cagttcatct taagaatgtc 300
 aacgattagt catgcaataa atgttctggg tttaaaaaat acatatctgg ttttggttaag 360
 gtatttttaa tcaattaggc ttgtagtatc agtgaaatac tgtaggttta gggactgggc 420
 tagcttcata tcagatttac ttgttaagtg actgttttgg aatgtttact tttggactgg 480
 gtttgtaaca cggttaaagg caatgagaaa caagcagaat tccaggagtc cttgaagcag 540
 agggcactgg aagacaatat agca 564

<210> 91
 <211> 540
 <212> DNA
 <213> Homo sapien

<400> 91
 atgcatgccg agcggcgagc tgtgatggat cgtggtcgag gccgaggtcg gcctaccacc 60
 ttcagaagtc gacctgtggc aaatgtggct accctgcaa gcgcaagaga aagtataact 120
 ggagtgcgaa ggctaaaaga cgaaatacca ccggaactgg tcgaatgagg cacctaaaaa 180
 ttgtataccg cagattcagg catggattcc gtgaaggaa aacacctaaa cccaagaggg 240
 cagctgttgc agcatccagt tcatcttaag aatgtcaacg attagtcatg caataaatgt 300
 tctgggttta aaaaatacat atctgggttt ggtaaggat ttttaataca ttaggcttgt 360
 agtatcagtg aaatactgta ggtttaggga ctgggctagc ttcatatcag atttacttgt 420
 taagtgactg ttttggaatg ttacttttg gactgggttt gtaacacggg taaaggcaat 480
 gagaaacaag cagaattcca ggagtccttg aagcagaggg cactggaaga caatatagca 540

135

<210> 92
 <211> 1103
 <212> DNA
 <213> Homo sapien

<400> 92
 gatgcatgct cgagcggcgc aggtgatgga tgcgcccggg caggtggccg agcaggaggc 60
 gccatcatgg gaggggacat ccgccataac aaggaccgaa aggttcggcg caaggagccc 120
 aagagccagg atatctacct gaggctgttg gtcaagttat acaggtttct ggccagaaga 180
 accaactcca cattcaacca ggttggtgtg aagaggttgt ttatgagtcg caccaaccgg 240
 ccgcctctgt ccctttcccg gatgatccgg aagatgaagc ttcctggccg ggaaaacaag 300
 acggccgtgg ttgtggggac cataactgat gatgtgcggg ttcaggaggt acccaaactg 360
 aaggtatgtg cactgcgcgt gaccagccgg gccgcagcc gcctcctcag ggccgggggc 420
 aagatcctca ctttcgacca gctggccctg gactccctta agggctgtgg cactgtcctg 480
 ctctccggtc ctcgcaaggg ccgagaggtg taccggcatt tcggcaaggc ccaggaacc 540
 ccgcacagcc acaccaagtg agtatcaggc cccagccct gccctctccc cagactcagc 600
 ctgcagggcc aggccctggc aacttgggc tgcttctcct atccctcgta ccagccagcc 660
 ccaggccttc tggaactctc agctgcctca tcctgcctgc ttttcccttc aactgtctt 720
 ggagctgtgt tttccagggc ctgagcaggc atctcctggg gagcatgcag ttcaccggcc 780
 tttgccagcc caggctggga ccctccaggc ggtggggcag gtgcttccat agccctcct 840
 ggctccttgg cttccactc tgccaagctt tccagcagc attgactctc ctggggctga 900
 atgtgacagg cttgtggtgt tcctgtgcca ctccacaatg gcctccttta agagggcccc 960
 cctctcactc cctcctctcc ccacagacc tacgtccgct ccaagggccg gaagtccgag 1020
 cgtgccagag gccgacgggc cagccgaggc tacaaaaact aaccctggat cctactctct 1080
 tattaataag atttttgctg aca 1103

<210> 93
 <211> 398
 <212> DNA
 <213> Homo sapien

<400> 93
 gatgcatgct cgagcggcgc aggtgatgga tgcgcccggg caggtggccg agcaggaggc 60
 gccatcatgg gaggggacat ccgccataac aaggaccgaa aggttcggcg caaggagccc 120
 aagagccagg atatctacct gaggctgttg gtcaagttat acaggtttct ggccagaaga 180
 accaactcca cattcaacca ggttggtgtg aagaggttgt ttatgagtcg caccaaccgg 240
 ccgcctctgt ccctttcccg gatgatccgg aagatgaagc ttcctggccg ggaaaacaag 300

136

acggccgtgg ttgtggggac cataactgat gatgtgcggg ttcaggaggt acccaaactg 360
atcggccgcg accacgctaa gccggattcc agcacact 398

<210> 94
<211> 1673
<212> DNA
<213> Homo sapien

<400> 94
gccagtgtga tggattggtc gcggcgaggt ggccgagcag gaggcgccat catgggagtg 60
gacatccgcc ataacaagga ccgaaagggt cggcgcaagg agcccaagag ccaggatatc 120
tacctgaggc tgttgggtcaa gttatacagg tttctggcca gaagaaccaa ctccacattc 180
aaccagggtg tgttgaagag gttgtttatg agtcgcacca accggccgcc tctgtccctt 240
tcccggatgg tgagtggctg gtccagagag caggttagac ctgggagccg ctgggtcttg 300
tctgtctgga agggaggcag gaacttgagc tctgggagca accagggcat caagggtttg 360
tctcagccag tcgcgagcgt agagcttgga ctgttggttg gaactgagtg tccatgggca 420
gtgggcaagt ctcttggtcc tcctctcctt ctctttgtga ggtggaggtg tcctgggtggc 480
ttcaggagat tgccccaagt cataactgag ttttatgtca aaggctctgc tgagggaggc 540
cctattgaac agtctgcatt ttttttttta tctggagctt tcccagctg gacctgagga 600
gaggcctgtg ggggtttaagg ggtgtggtgt gagtgggggg tcccgtaga gggcacagct 660
gccctggcct ggggatgcct caagcggggc tgaatgtaaa caccagaaca acttacgacg 720
tacatcctcc ccaccctaag atccggaaga tgaagcttcc tggccgggaa aacaagacgg 780
ccgtgggtgt ggggaccata actgatgatg tgcgggttca ggaggtagcc aaactgaagg 840
tgagctggcg ggggctgggc agaccatca gacccttgct gtactgtgct gtgctgtgct 900
gtgctgtgct gtgctgtgct ggtccatcct cagtcttgcc accgctggtg aggcaagcat 960
cctgcctgtc tgagtggctt tgaggattgg tgtgggtcgt ctctggcatg gttgatggtg 1020
gggtcccctg ggcaagcttc ttctggggca agttgtaaac tgtagtgct ggcccctaga 1080
gccagacagc ctgggctcaa gtctagactc ctttgtcacc tctgcacagg tctttttttg 1140
gccagcttc cttgctgtag tttctagcat aacaccaggg atgtggtggt gaaagttccg 1200
gaagtgttga gaactgctgc tgtcagccca ggggtgctccg ggtagtaaac gatagtgttg 1260
gttaccaggg ctctgctgcg ccaaggcagt ggagggtagg tcctggcagc ccctgggtga 1320
ccacgcaggg gtctgacctg ccagccttgt cctctccacc aggtatgtgc actgcgcgtg 1380
accagccggg cccgcagccg catcctcagg gcagggggca agatcctcac tttcgaccag 1440
ctggccctgg actcccctaa gggctgtggc actgtcctgc tctccggtcc tcgcaagggc 1500

137

cgagaggtgt accggcattt eggcaaggcc ccaggaaccc cgcacagcca caccaaacc 1560
 tacgtccgct ccaagggccg gaagttcgag cgtgccagag gccgacgggc cagccgaggc 1620
 tacaaaaact aaccctggat cctactctct tattaaaaag atttttgctg aca 1673

<210> 95
 <211> 1121
 <212> DNA
 <213> Homo sapien

<400> 95
 tgagtgcctg gaccctgacc tccaacggga gatccttccc ggggccattc cctaaatcgt 60
 gggcttgctt cttgctccct ctggctattc tgtgtccctg tggctgctcc ccaaccctaa 120
 gggctgtccc tgatctctct tgetatttcc ccaagcccgga gactgctccc ttacaaattc 180
 tggctgctcc ctttccctgt gtccaatacc gtgtgatcac ctgtccctcc ctggctccctc 240
 taattctccc ttgtgactat tctgtaatcc cactccccgt tcctgagccc cggggattgt 300
 ttcttggttc ccagaaatgt tctccaagaa ccaaagtgc tgttccctaa tccactgac 360
 ccttcctcag cacatctgct tgttgctctg cagctctgac tattccagac accagaggggt 420
 gttacactga cctcctggct agtccaagcc tgttccctga tggctcctga tggcctctga 480
 gccagtgac tatcttgaag ccccaactggc tattcctgtg cttaaactcc ttagccgcca 540
 accgggtggg gcccttact actgccagtt tttccctgtc ctttgggaca attccacagg 600
 cgacctgcct gactgacccc cagggcctat tccccagcca ccattccagg aactgattt 660
 actggcttac tgttaccttt caggacccag tgacatttcc ctcaatggga ggggctgtcc 720
 tggggtatgt acttaggtgg ggtgagagca agtctcccca gcggctcagt gttgaaagct 780
 gaggagagga agttctgctc cacggggggag gcaggggagaa gccctcagag atgtttccct 840
 cctcccttcc aggtgtttcc ctgtttatcc atcccctgca aactgcagag tggcactcat 900
 tgcttggtga cggaccagct cctccaaggc tctgaaaagg gcttccagtc ccgtcacctt 960
 gcctgcctgc ctggggagcc agggctgtgc acctggcagt ccctgcggtc ccagatagcc 1020
 tgaatcctgc ccggagtga agctgaagcc tgcacagtgt ccaccctgtt cccactccca 1080
 tctttcttcc ggacaatgaa ataaagagtt accaccagc a 1121

<210> 96
 <211> 570
 <212> DNA
 <213> Homo sapien

<400> 96
 ccagtccttg gggcgcggcc tagcaaccac gagggggcga ggctctgatg ggaatgggtcc 60

138

cactggaaat	ggggaccacc	ccaatttcag	tctatcagaa	ggcggggcct	ttggcagctt	120
ggcagctcag	ccaatcacia	gctgcttgct	ggcccccgcc	ccgcctttct	ccttaggaaa	180
actgcaggat	ggactcttgc	acatcactac	ctgcagtttt	gtggctccct	ggaacagcct	240
gagcttagct	cagcgccggg	gcttcaccaa	gacctacact	gttggtgtg	aggaatgcac	300
agtgtttccc	tgtttatcca	tcccctgcaa	actgcagagt	ggcactcatt	gcttgtggac	360
ggaccagctc	ctccaaggct	ctgaaaaggg	cttcagctcc	cgtcaccttg	cctgcctgcc	420
tcgggagcca	gggctgtgca	cctggcagtc	cctgcgggtcc	cagatagcct	gaatcctgcc	480
cggagtggaa	gctgaagcct	gcacagtgtc	caccctgttc	ccactcccat	ctttcttccg	540
gacaatgaaa	taaagagtta	ccaccagca				570

<210> 97

<211> 4990

<212> DNA

<213> Homo sapien

<400> 97

gggagtcggc	ggcacaaaat	ggcggcgggc	gcggcgggcg	ctgggtgctgc	agggtcggca	60
gctccccggg	cagcgggccgg	cgccccggga	tctggggggc	caccctcagg	gtcgcagggg	120
gtgctgatcg	gggacaggct	gtactccggg	gtgctcatca	ccttgagaaa	ctgcctcctg	180
cctgacgaca	agctccgttt	cacgccgtcc	atgtcgagcg	gcctcgacac	cgacacagag	240
accgacctcc	gcgtgggtgg	ctgcgagctc	atccaggcgg	ccggtatcct	gctccgcctg	300
ccgcagggtg	ccatggctac	cgggcagggt	ttgttccagc	ggttctttta	taccaagtcc	360
ttcgtgaagc	actccatgga	gcatgtgtca	atggcctgtg	tccacctggc	ttccaagata	420
gaagaggccc	caagacgcat	acgggacgtc	atcaatgtgt	ttcaccgcct	tcgacagctg	480
agagacaaaa	agaagcccgt	gcctctacta	ctggatcaag	attatgttaa	tttaaagaac	540
caaattataa	aggcggaag	acgagttctc	aaagagttgg	gtttctgcgt	ccatgtgaag	600
catcctcata	agataatcgt	tatgtacctt	cagggtgttag	agtgtgagcg	taaccaacac	660
ctgggtccaga	cctcatgggt	agcctctgag	ggtaagtgc	taagacttct	cctctgctgt	720
ccaagcgctt	tggtgcaggg	acagcggcat	cttcagccaa	tccagtgcag	gctctccacc	780
gaaggctggc	tctagactgg	tggtacgcac	atagcatagc	catggccgac	tcctgctgtg	840
gttctctgac	gattgtgctt	cttggttaatc	ctctgtcgtg	ctttggtaat	cgtattgatt	900
agagttggta	actgtcttga	cttgaatttt	gtccctttta	aactgctgta	cctgtatgat	960
aaagatgcag	tacctttctc	ttaaaaaaaa	atgctatgga	aagctgtgag	aattgaagag	1020
acaaattggc	tgtgtcagtg	tggggttatg	tcatgatttc	tagaagccct	gaagttgctc	1080

139

ttttgagcag ctttgcacga cacgctctgg taaaagggtg gcatctttaa attatttcat	1140
ggatactttg aaaaatattg tatcacttca aatacagcaa taagtttata tggtctcaag	1200
atttcatttg tttttaagaa ttttaagttc gtggattaat atcactactt gaatactgac	1260
agttgttgat tagacaccga aaggttactg attgttgaat gtatctgtgt tagagctgtg	1320
cactggcacg cttgcatcag gggctggggc cacacggccg ccacacagat tccccctga	1380
tgcttgagc tgcttccaga gccgggtgtc tccaagaggc acctgtagga cttcccat	1440
agaaatctct tgagtgggtt tgtatgttac cttctccaag gtttatttag gacagagata	1500
ttgttggaag gtcacgggtc agattccctc acaaccacc tcgtctgcgg gtgcagcccc	1560
actccaaggc tccccgttat tggggtatgt gaggagcagt aaatataaaa ccagttcaac	1620
tgctctcatg gaatcaccct ttctgttttt gcagtattca taaagctagt gtaaggctctg	1680
gttttagtct attaaatctt agagatctaa aggaaatgct caaaatgtag ccaggtttta	1740
aatgctttaa cttttaaaaa atgtaaattt ttgtatgttt atagcttcta aatatgaaag	1800
ttaaagaatg tactgtgatg aaatgttcag tattatgttg cttctcagta tcatgttgct	1860
tctcagtatt gtgttgcttc tgattctatg aatgttcatt ttaagacccc ttgttgaaat	1920
gggacagtgt gcagcggctc tgatgagccc gagaagaggc ctgcccttgg gtgcggagtc	1980
tccctccgca cgatgctccc acgctccaa cttgcaccca aggggctttt ccctcttcca	2040
agtggactcc ttcaaggaag ctgcagctcg gtcagcagag aaggggcctg ccgccagcgc	2100
cctggaggaa gaggaagagg aaccaagag gatggcttgt ctcccagcag ccacaccggc	2160
tttgtgtcga gccagttcat ttgagtttgc atgtttctct gcactatgga ttttgagcat	2220
ttagatttct ttaatcaaaa gcgttttagt gactccagta gacattttct ttctgaggca	2280
tcgtgctttg catgagagca ggccaagggt gaggggaaaa gtaaagttaa agtcggttct	2340
ctttcatagc aacacgtatt gtctgacatt cagccagctt ttttttttcc taataatttc	2400
tgtgcctttc tgtcctgtat ttactgtatt tagaaaaagc agctagaata tttctccatt	2460
aactcttgag attcacagga ctgtctagct ctgagtccta gcaatagact ccttagagga	2520
gtagtacgtt tatctagatt ttctctagat aatgcaggcg gaagacctgg gttcccgggt	2580
ggggcattgc agttcttcct gtgtttggct tccaggaatt acatgaacga cagccttcgc	2640
accgacgtct tcgtgcgggt ccagccagag agcatcgct gtgcctgcat ttatcttgct	2700
gcccggacgc tggagatccc tttgccaat cgtccccatt ggtttctttt gtttgagca	2760
actgaagaag aaattcagga aatctgctta aagatcttgc agctttatgc tcggaaaaag	2820
gttgatctca cacacctgga ggggtgaagt gaaaaaagaa agcacgctat cgaagaggca	2880
aaggcccaag cccggggcct gttgcctggg ggcacacagg tgctggatgg tacctcgggg	2940

140

ttctctcctg	cccccaagct	ggtggaatcc	cccaaagaag	gtaaagggag	caagccttcc	3000
ccactgtctg	tgaagaacac	caagaggagg	ctggagggcg	ccaagaaagc	caaggcggac	3060
agccccgtga	acggcttgcc	aaaggggcga	gagagtcgga	gtcggagccg	gagccgtgag	3120
cagagctact	cgagggtcccc	atccccgatca	gcgtctccta	agaggaggaa	aagtgcacagc	3180
ggctccacat	ctggtgggtc	caagtcgcag	cgctccctac	aaaggctctg	agattcgggg	3240
ctcccgggaag	tccaaggact	gcaagtaccc	ccagaagcca	cacaagtctc	ggagccggag	3300
ttcttcccgt	tctogaagca	ggtcacggga	gcgggcggat	aatccgggaa	aatacaagaa	3360
gaaaagtcac	tactacagag	atcagcgacg	agagcgctcg	aggtcgtatg	aacgcacagg	3420
ccgtcgctat	gagcgggacc	accctgggca	cagcaggcat	cggagggtgag	gcgggggttgc	3480
agtgactggg	ggccgcaagc	ccttccctgg	ggagtaacctg	atggctgccc	tttgaccccc	3540
ggtggctgcc	ctttgacccc	cgggtgtgct	ctcagcgcaa	gtggctcctag	aacaggattc	3600
tttttgaaaa	tgtctgtcga	ctggaccttg	gtggatttgg	aaatggaact	gagggaccgg	3660
tgacacgtgc	ttcagaccgg	tctgggggtgc	ggcgcacacc	tgggcccgtg	cagggctcag	3720
ctcggcagca	gctctgaggg	cagctcaatg	aaaaagtga	tgcacacgcc	cttggtggcg	3780
tggcctggca	tggcctgggtg	ctatcggcag	ccgctctcca	ctccccgact	gataactcaat	3840
tacgtgaagc	caagaaagat	gattttttaga	acctttgcct	atattagggt	gtacttatgt	3900
acatatTTTTg	cagtgtttca	caggagaaag	tggccttaac	tgccccttat	tctctctcca	3960
cgttgtaaat	aaacatgtgt	ttaatacaag	ttaaagctat	gtatgaaaac	tcagaacttg	4020
aatcccgtca	gcttaaaaact	tgtgtaggga	atcctgactt	ttaaaatgtg	agggtatttg	4080
gatctgtgtt	gaaagtcgta	tatttttatc	tgtgcgggtgc	tgagtgcagg	ccaccagctc	4140
ctaaatagag	gttccctata	tgcgcgtatg	acatggtgaa	taaacacaac	tctctccact	4200
caggacatcc	ggagcgttat	ggacgtggta	ggtggctcgtt	ctgtgtgctt	gtgaaagtgt	4260
ccaggcgtgt	gcacagccag	tgcgcccact	tccgggctcc	ttgctccctg	ctgtactgaa	4320
gttttggaatt	ttgcatccaa	tcctgtgtgc	ctgcccttct	gccgaaggct	tgtgaggggc	4380
ctgagtcctc	tgcccatcag	gatgacaggc	tccttcctgc	agggccatag	gagggaaagt	4440
ttggaaacac	agaatgattc	caagtgctc	tcgttcctga	gggggactgg	tttgtaaccc	4500
atgacatctg	tgggcgagag	aggcagctgg	gagcaggaca	cttggagggt	cacccacagg	4560
gggtggcacc	tgactctga	gtgccccca	ctgtcatcag	ctgcctctta	ccgtggacac	4620
agttttgggt	ttggggacta	gggggccccca	ctcctgggtg	taccgttttg	acttactagg	4680
gcagtgggac	atataggccg	gggctagtgg	gataacgggg	agttacgcct	gatgactttt	4740

141

ttgatggaat cctgcattag atagctggtg ggaccccccc ctcagaattg gggaaactgag	4800
gagactccag ggagggtgtc cttccagggg gagcagctat gaggggcccc ctagcttcct	4860
gtgcctggaa gtaagagaac cagtaaaggg ccatacacac ctgtacccaa gagaccgctc	4920
tccatttgct ttcttttttt actaaataat tgtaaaatat tattatgaca taaagaacca	4980
tttaaggcca	4990

<210> 98

<211> 4944

<212> DNA

<213> Homo sapien

<400> 98

gggagtcggc ggcacaaaat ggcggcgggcg gcggcgggcg ctggtgctgc agggtcggca	60
gctcccgcg cagcgcccg cgccccggga tctgggggcg caccctcagg gtgcgagggg	120
gtgctgatcg gggacaggct gtactccggg gtgctcatca ccttgagaaa ctgcctcctg	180
cctgacgaca agctccggtt cagcccgctc atgtcgagcg gcctcgacac cgacacagag	240
accgacctcc gcgtgggtgg ctgcgagctc atccaggcg cggtatcct gctccgctg	300
ccgcagggtg ccatggctac cgggcagggt ttgttccagc gggtctttta taccaagtcc	360
ttcgtgaagc actccatgga gcatgtgtca atggcctgtg tccacctggc ttccaagata	420
gaagaggccc caagacgcat acgggacgtc atcaatgtgt ttcaccgcct tcgacagctg	480
agagacaaaa agaagcccgt gcctctacta ctggatcaag attatgttaa tttaaagaac	540
caaattataa aggcggaaag acgagttctc aaagagttgg gtttctgcgt ccatgtgaag	600
catcctcata agataatcgt tatgtacctt cagggtgttag agtgtgagcg taaccaacac	660
ctggtccaga cctcatgggt agcctctgag ggtaagtgtc taagacttct cctctgctgt	720
ccaagcgctt tgggtgcagg acagcggcat cttcagccaa tccagtgcag gctctccacc	780
gaaggctggc tctagactgg tggtagcgac atagcatagc catggccgac tcctgctgtg	840
gttctctgac gattgtgctt cttgttaatc ctctgtcgtg ctttggtaat cgtattgatt	900
agagttggta actgtcttga cttgaatttt gtccctttaa aactgctgta cctgtatgat	960
aaagatgcag tacctttctc ttaaaaaaaaa atgctatgga aagctgtgag aattgaagag	1020
acaaattggc tgtgtcagtg tggggttatg tcatgatttc tagaagccct gaagttgctc	1080
ttttgagcag ctttgcatga cagctctgg taaaagggtg gcatctttaa attatttcat	1140
ggatactttg aaaaatattg tatcacttca aatacagcaa taagtttata tgttctcaag	1200
atttcatttg tttttaagaa ttttaagttc gtggattaat atcactactt gaatactgac	1260
agttgttgat tagacaccga aaggttactg attgttgaat gtatctgtgt tagagctgtg	1320

142

cactggcacg cttgcatcag gggctggggc cacacggccg ccacacagat tcccccgatga	1380
tgccctggagc tgcttccaga gccgggtgtc tccaagaggc acctgtagga cttcccat	1440
agaaatctct tgagtgggtt tgtatgttac cttctccaag gtttatttag gacagagata	1500
ttgctggaag gtcattgggtc agattccctc acaaccacc tcgtctgcgg gtgcagcccc	1560
actccaaggc tccccgttat tgggggtatgt gaggagcagt aaatataaaa ccagttcaac	1620
tgctctcatg gaatcaccct ttctgttttt gcagtattca taaagctagt gtaaggctctg	1680
gttttagtct attaaatctt agagatctaa aggaaatgct caaaatgtag ccaggtttta	1740
aatgctttaa cttttaaaaa atgtaaat	1800
ttatgtgtt atagcttcta aatatgaaag	1860
ttaaagaatg tactgtgatg aaatgttcag tattatgttg cttctcagta tcatgttgct	1860
tctcagtatt gtgttgcttc tgattctatg aatgttcatt ttaagacccc ttgttgaaat	1920
gggacagttg gcagcggctc tgatgagccc gagaagaggc ctgcccttgg gtgcggagtc	1980
tccctccgca cgatgctccc acgcgtccaa cttgcaccca aggggctttt ccctcttcca	2040
agtggactcc ttcaaggaag ctgcagctcg gtcagcagag aaggggcctg ccgccagcgc	2100
cctggaggaa gaggaagagg aacccaagag gatggcttgt ctcccagcag ccacaccggc	2160
tttgtgctca gccagttcat ttgagtttg atgtttctct gcactatgga ttttgagcat	2220
ttagatttct ttaatcaaaa gcgttttagt gactccagta gacattttct ttctgaggca	2280
tcgtgctttg catgagagca ggccaagggt gaggggaaaa gtaaagttaa agtcggttct	2340
ctttcatagc aacacgtatt gtctgacatt cagccagctt tttttttt taataatttc	2400
tgtgcctttc tgtcctgtat ttactgtatt tagaaaaagc agctagaata tttctccatt	2460
aactcttgag attcacagga ctgtctagct ctgagtccta gcaatagact ccttagagga	2520
gtagtacgtt tatctagatt ttctctagat aatgcaggcg gaagacctgg gttcccgggt	2580
ggggcattgc agttcttctt gtgtttggct tccaggaatt acatgaacga cagccttcgc	2640
accgacgtct tcgtgcgggt ccagccagag agcatcgctt gtgcctgcat ttatcttgct	2700
gcccggacgc tggagatccc tttgccaat cgtccccatt ggtttctttt gtttgagca	2760
actgaagaag aaattcagga aatctgctta aagatcttgc agctttatgc tcggaaaaag	2820
gttgatctca cacacctgga ggggtgaagt gaaaaaagaa agcacgctat cgaagaggca	2880
aaggcccaag cccggggcct gttgcctggg ggcacacagg tgctggatgg tacctcgggg	2940
ttctctcctg cccccaagct ggtggaatcc cccaaagaag gtaaaggag caagccttcc	3000
ccactgtctg tgaagaacac caagaggagg ctggaggggc ccaagaaagc caaggcggac	3060
agccccgtga acggcttgcc aaagggcgga gagagtcgga gtcggagccg gagccgtgag	3120
cagagctact cgagggtccc atcccgatca gcgtctccta agaggaggaa aagtgcagc	3180

143

ggctccacat	ctggtgggtc	caagtgcgag	agccgctccc	ggagcaggag	tgactcccca	3240
ccgagacagg	ccccccgcag	cgctccctac	aaaggctctg	agattcgggg	ctcccgggaag	3300
tccaaggact	gcaagtaccc	ccagaagcca	cacaagtctc	ggagccggag	ttcttcccgt	3360
tctcgaagca	ggtcacggga	gcgggcggat	aatccgggaa	aatacaagaa	gaaaagtcac	3420
tactacagag	atcagcgacg	agagcgctcg	aggctgatat	aacgcacagg	ccgtcgctat	3480
gagcgggacc	acccggtggc	tgccctttga	ccccgggtg	tgctctcagc	gcaagtggtc	3540
ctagaacagg	attctttttg	gaaatgtctg	tcgactggac	cttggtggtg	ttggaaatgg	3600
aactgagggg	ccggtgacac	gtgcttcaga	ccggtctggg	gtgcggcgca	cacctggggc	3660
cgtgcagggc	tcagctcggc	agcagctctg	agggcagctc	aatgaaaaag	tgaatgcaca	3720
cgcccttggt	ggcgtggcct	ggcatggcct	ggtgctatcg	gcagccgctc	tccactcccc	3780
gactgatact	caattacgtg	aagccaagaa	agatgatttt	tagaaccttt	gcctatatatta	3840
ggttgtactt	atgtacatat	tttgagtggt	ttcacaggag	aaagtggcct	taactgcccc	3900
ttattctctc	tccacgttgt	aaataaacat	gtgtttaata	caagttaaag	ctatgtatga	3960
aaactcagaa	cttgaatccc	gtcagcttaa	aacttggtga	gggaatcctg	acttttaaaa	4020
tgtgagggta	tttgatctg	tggtgaaagt	cgtatatatt	tatctgtgcg	gtgctgagtg	4080
cagggcacca	gctcctaaat	agaggttccc	tatatgcgcg	tatgacatgg	tgaataaaca	4140
caactctctc	cactcaggac	atccggagcg	ttatggacgt	ggtaggtggg	cgttctgtgt	4200
gcttgtagaa	gtgtccaggc	gtgtgcacag	ccagtgcgcc	cacttcgggg	ctccttgctc	4260
cctgctgtac	tgaagttttg	gattttgcat	ccaatcctgt	gtgcctgccc	ttctgccgaa	4320
ggcttgtagg	gggcctgagt	cctctgcccc	tcaggatgac	aggctccttc	ctgcagggcc	4380
ataggagggg	agttttggaa	acacagaatg	attccaaggt	gctctcgttc	ctgagggggg	4440
ctggtttgta	acccatgaca	tctgtgggcg	agagaggcag	ctgggagcag	gacacttgga	4500
gggtcacccc	acgggggtgg	cacctgcact	ctgagtgcc	ccactgtca	tcagctgcct	4560
cttaccgtgg	acacagtttt	ggttttgggg	actagggggc	cccactcctg	gtggtaccgt	4620
ttggacttac	tagggcagtg	ggacatatag	gccggggcta	gtgggataac	ggggagttac	4680
gcctgatgac	ttttttgatg	gaatcctgca	ttagatagct	ggtgggaccc	ccccctcaga	4740
attggggaac	tgaggagact	ccagggaggg	tgctcttcca	gggagagcag	ctatgagggg	4800
ccccctagct	tcctgtgcct	ggaagtaaga	gaaccagtaa	agggccatac	acacctgtac	4860
ccaagagacc	gctctccatt	tgctttcttt	ttttactaaa	taattgtaaa	atattattat	4920
gacataaaga	accatttaag	gcca				4944

144

<210> 99
 <211> 438
 <212> DNA
 <213> Homo sapien

<400> 99
 gcctactact actaaattcg cggccgcgctc gaccttcttg aactgggtgct gtctgggctt 60
 catagcattc gcctactccg tgaagtctag ggacaggaag atggttggcg acgtgaccgg 120
 ggcccaggcc tatgcctcca ccgccaagtg cctgaacatc tgggccctga ttttgggcat 180
 cttcatgacc ggattcatcc tgttactggg attcggctct gtgacagtct accatattat 240
 gttacagata atacaggaaa aacgggggta ctagtagccg cccatagcct gcaacctttg 300
 cactccactg tgcaatgctg gccctgcacg ctggggctgt tgcccctgcc cccttggtcc 360
 tgcccctaga tacagcagtt tatacccaca cacctgtcta cagtgtcatt caataaagtg 420
 cacgtgcttg tgaaaaar 438

<210> 100
 <211> 450
 <212> DNA
 <213> Homo sapien

<400> 100
 agattaatat atagggcaat tggtcactca atcatgctcg agcggcgagc tgtgatggat 60
 yggycgcccgg scgagggtacg tgaaacccaa attaagggta aaagtgggag attcttcaca 120
 gtcaagctcc cagttgctct tgatcctggg gccaaagattt cagtcattgt ggaaacagtc 180
 tacacccatg tgcttcatcc atatccaacc cagatcaccc agtcagagaa acagtttgtg 240
 gtgtttgagg ggaaccatta tttctactct ccctatccaa cgaagacaca aaccatgctg 300
 gtgaagcttg cctctcgaaa tgtggagagc tacaccaagc tggggaaccc cacgcgctct 360
 gaggacctac tggattatgg gcctttcaga gatgtgcctg cctatagtca ggatactttt 420
 aaagtacctc ggccgcgacc acgctaagcc 450

<210> 101
 <211> 1965
 <212> DNA
 <213> Homo sapien

<400> 101
 agcggaatga acggctgctc agtggcgagg gggcgggcta ccgcgcccgg gccctacccg 60
 ccctggcgca ctgctcttcc cggtcattga ggcgccagcc gccggcttgt ttctgctcct 120
 gttgcttggg acttgggccc cggcgccggg cagcgccctc tccgaggcac cggcgctgat 180
 caatgaggac gtgaagcgca cagtggacct aagcagccac ctggctaagg tgacggccga 240

145

```

ggtaggtcctg ggcacactgg gcggcggtc cactgccga gctacctctt tcctgctggc 300
tttgagacct gagctcgagg cccggctggc gcacctgggc gtgcaggtaa agggagaaga 360
tgaggaagag aacaatttgg aagtacgtga aacaaaaatt aagggtaaaa gtgggagatt 420
cttcacagtc aagctcccag ttgctcttga tcctggggcc aagatttcag tcattgtgga 480
aacagtctac acccatgtgc ttcattcata tccaaccag atcaccagt cagagaaaca 540
gtttgtggtg tttgagggga accattatct ctactctccc tatccaacga agacacaaac 600
catgcgtgtg aagcttgccct ctcgaaatgt ggagagctac accaagctgg ggaacccac 660
gcgctctgag gacctactgg attatgggccc tttcagagat gtgcctgcct atagtcagga 720
tactttttaa gtacattatg agaacaacag ccctttcctg accatcacca gcatgaccgc 780
agtcattgaa gtctctcact ggggtaatat tgctgtggaa gaaaatgtgg acttaaagca 840
cacaggagct gtgcttaagg ggcctttctc acgctatgat taccagagac agccagatag 900
tggaatatcc tccatccgtt cttttaagac catccttcct gctgctgccc aggatgttta 960
ttaccgggat gagattggca atgtttctac cagccacctc cttatttttg atgactctgt 1020
agagatggaa atccggcctc gcttccctct ctttggcggtg tggaagaccc attacatcgt 1080
tggtacaac ctcccaagct atgagtacct ctataatttg ggtgaccagt atgcactgaa 1140
gatgaggttt gtggaccatg tgtttgatga acaagtata gattctctga ctgtgaagat 1200
catcctgcct gaaggagcca agaacattga aattgatagt ccctatgaaa tcagccgtgc 1260
cccagatgag ctgcactaca cctatctgga cacatttggc cgccctgtga ttgttgccca 1320
caagaaaaat ctggtagaac agcacattca ggacattgtg ctggatgcac aggtcaagga 1380
gctggtgctg aagtcggcgg tggaggctga gcgcctggtg gctggcaagc tcaagaaaga 1440
cactacatt gagaatgaga agctcatctc aggaagcgc caggagctgg tcaccaagat 1500
cgaccacatc ctggatgccc tgtagcccct gcccgcatcc tccagggggc ccagggtgcc 1560
tgcactttgc tgtggcaggc agattgggtg gtagtgggag gttgtgcatg gaggccagtg 1620
aaagctgaca tctgtaaaag gccttcaagg aagagaaacc aggccctgcg tcaggcagtg 1680
tgagtttgcc gtttgcctt aactttcttt tttttttttt ttaaaaaaag aaaacttta 1740
aaaaactccc attaaaaaca aaacatcttt gtgttttgaa caaaggaatt ttcaatattt 1800
gattgggtatt ctgttctgaa gtctaggata tttttcagcc tataaagccc cctgttttat 1860
gcccttctaa ttctgatgtt tgggtattgt gtgagtgcac gtgttttttt tttttttttt 1920
taaagcgtgt gtgaacaaat ggaaataaag cagggactgt gaaca 1965

```

<210> 102

<211> 1193

146

<212> DNA

<213> Homo sapien

<400> 102

```

agcgggaatga acggctgctc agtggcggga gggcgggcta ccgcgcccgg gccctaccgg      60
ccctggcgca ctgctcttcc cggtcattga ggcgccagcc gccggcttgt ttctgctcct      120
gttgcttggg acttggggccc cggcgccggg cagcgcctcc tccgaggcac cgccgctgat      180
caatgaggac gtgaagcgca cagtggacct aagcagccac ctggctaagg tgacggccga      240
ggcggctcctg gcgcacctgg gcggcggtc cagtcctcga gctacctctt tcctgctggc      300
tttgaggcct gagctcgagg cccggctggc gcacctgggc gtgcaggtaa agggagaaga      360
tgaggaagag aacaatttgg aagtacgtga aacaaaaatt aagggtaaaa gtgggagatt      420
cttcacagtc aagctcccag ttgctcttga tcctggggcc aagatttcag tcattgtgga      480
aacagtctac acccatgtgc ttcattcata tccaaccag atcaccagc cagagaaaca      540
gtttgtggtg tttgagggga accattatct ctactctccc tatccaacga agacacaaac      600
catgcgtgtg aagcttgctt ctcgaaatgt ggagagctac accaagctgg ggaacccac      660
gcgctctgag gacctactgg attatgggccc ttccagagat gtgcctgcct atagtcagga      720
tactttttaa gtacattatg agaacaacag ccctttcctg accatcacca gcatgacctg      780
agtcattgaa gtctctcact ggggtaatat tgctgtggaa gaaaatgtgg acttaaagca      840
cacaggagct gtgcttaagg ggcctttctc acgctatgat taccagagac agccagatag      900
tggaatatcc tccatccgtt cttttaagac catccttctt gctgctgccc aggatgttta      960
ttaccgggat gagattggca atgtttctac cagccacctc cttatttttg atgactctgt     1020
agagatggaa atccggcctc gcttccctct ctttggcggg tggaagacc attacatcgt     1080
tggtctacaac ctcccaagct atgagtacct ctataatttg ggtcagtcct caatagtaag     1140
ggagaaacta accttttctt taatttcgtg aatttttttt ctttcaatta tca             1193

```

<210> 103

<211> 1685

<212> DNA

<213> Homo sapien

<400> 103

```

agcgggaatga acggctgctc agtggcggga gggcgggcta ccgcgcccgg gccctaccgg      60
ccctggcgca ctgctcttcc cggtcattga ggcgccagcc gccggcttgt ttctgctcct      120
gttgcttggg acttggggccc cggcgccggg cagcgcctcc tccgaggcac cgccgctgat      180
caatgaggac gtgaagcgca cagtggacct aagcagccac ctggctaagg tgacggccga      240
ggcggctcctg gcgcacctgg gcggcggtc cagtcctcga gctacctctt tcctgctggc      300

```

147

```

tttggagcct gagctcgagg cccggctggc gcacctgggc gtgcaggtaa agggagaaga      360
tgaggaagag aacaatttgg aagtacgtga aaccaaaatt aagggtaaaa gtgggagatt      420
cttcacagtc aagctcccag ttgctcttga tcctggggcc aagatttcag tcattgtgga      480
aacagtctac acccatgtgc ttcattccata tccaaccag atcaccagt cagagaaaca      540
gtttgtggtg tttgagggga accattattt ctactctccc tatccaacga agacacaaac      600
catgcgtgtg aagcttgccct ctcgaaatgt ggagagctac accaagctgg ggaacccac      660
gcgctctgag gacctactgg attatgggcc tttcagagat gtgcctgcct atagttagga      720
tactttttaa gtacattatg agaacaacag ccctttcctg accatcacca gcatgaccgc      780
agtcattgaa gtctctcact ggggtaatat tgctgtggaa gaaaatgtgg acttaaagca      840
cacaggagct gtgcttaagg ggcctttctc acgctatgat taccagagac agccagatag      900
tggaatatcc tccatccgtt cttttaagac catccttcct gctgctgccc aggatgttta      960
ttaccgggat gagattggca atgtttctac cagccacctc cttatttttg atgactctgt    1020
agagatggaa atccggcctc gcttccctct ctttggcggg tggaagacc attacatcgt    1080
tggctacaac ctcccaagct atgagtacct ctataatttg ggtgaccagt atgcactgaa    1140
gatgaggttt gtggaccatg tgtttgatga acaagtata gattctctga ctgtgaagat    1200
catcctgcct gaaggagcca agaccacatc ctggatgccc tgtagccct gcccgcatcc    1260
tccagggggc ccagggtgcc tgcactttgc tgtggcaggc agattgggtg gtagtgggag    1320
gttgtgcatg gaggccagtg aaagctgaca tctgtaaaag gccttcaagg aagagaaacc    1380
aggccctgcg tcaggcagtg tgagtttgcc gtttgcctt aactttcttt tttttttttt    1440
ttaaaaaaag aaaactttaa aaaaactccc attaaaaaca aaacatcttt gtgttttgaa    1500
caaaggaatt ttcaatatatt gattggtatt ctgttctgaa gtctaggata tttttcagcc    1560
tataaagccc cctgttttat gcccttctaa ttctgatgtt tgggtattgt gtgagtgcac    1620
gtgttttttt tttttttttt taaagcgtgt gtgaacaaat ggaaataaag cagggactgt    1680
gaaca                                              1685

```

```

<210> 104
<211> 1182
<212> DNA
<213> Homo sapien

```

```

<400> 104
gggttgcgcc agggagccgg aaagccgact cccgaagttg gggctctggg aaaacttggg      60
tcctgggttg actgagaagc ggcggggaaa ggaggcgggc caggaggagg gggcctggcg    120
gacgcgggcc ggggggcggg gcgcggcggg gctgtcggtc acgcccctca gtccgccccg    180

```

148

```

ccccgccccg cctgccgggg aagggccacg ttgcccggcc ggccgtccgg cccgggcgcg      240
ccgcagaaaag ggctggcgag tcgaaaggcg aggcggccgc ggcagcgctt gggacgcgcc      300
tgggcaccgg gctcgctccc tgcgccccgg agcaggccaa gttcggggcc aggacgtcgg      360
gaggacctgg tgcattggctg cctcctaata ccatagtcca gaggaggcat ccctaggact      420
gcgggcaagg gagccgggca agcccagggc agccttgaac cgtcccctgg cctgccctcc      480
ccggtggggg ccaggatgct gaagaagcag tctgcagggc ttgtgctgtg gggcgctatc      540
ctctttgtgg cctgatacta accaaagggg ctctctgtgc agggcctggg ggagttgggtg      600
ggtcacccgg gctcactgcc tctgcccctt ctctcctgtc tgacccccac ttagcccttc      660
tctccttgca gcctagcagt ttatagttct gagatggaaa gttgaagggg gcaagcaaga      720
cctctcctca gcccatgccc agctgtcagg agagagggtgc agggaggaag gccttgtgct      780
gggacaacct ctctcttgcc ttacctcaga gagggactat gcctgaccc ctctttctg      840
aaaatcagtg ccctccctgt tgctctagga ggctcctgct ggcttggtag aagacagaat      900
tcgatctgcc tgtccctttt tcccctgggg tttagacacac aggtcctct cagcatgagg      960
tggagcagtg accagggtga gcagtgacca ggacgcctct ggcccagtgc tgcccagcct     1020
ccccgcccgc tcccaggcgc cccatgtcct cacaggccag gacgccatgg caggatggag     1080
aggacttggg ggatttttgt ttcttgctg acctcagttt catgaaagaa agtggaagct     1140
acagaattat tttctaaaat aaaggctgaa ttgtctgaaa aa                        1182

```

<210> 105
 <211> 2645
 <212> DNA
 <213> Homo sapien

```

<400> 105
agcgggcgcg ctccgctcgg cagcctgtgg gacgcgaccg cggcgctagt ctcgttcctt      60
tgtgctgcgg cggcggcttc tcgagtcctc cccgacgcgt cctctaggcc agcgagcccc     120
gcgctctccg gtgacggacc atgtcggcgg cgggagcggg cgcgggcgta gaggcgggct     180
tctccagcga ggagctgctc tcgctccgtt tcccgtgca ccgcgcctgc cgcgacgggg     240
acctggccac gctctgctcg ctgctgcagc agacacccca cggccacctg gcctctgagg     300
actccttcta tggctggacg cccgtgcact gggccgcgca tttcggcaag ttggagtgtc     360
tagtgagttt ggtgagagcg ggagccacac tcaacgtctc caccacacgg tacgcgcaga     420
cgccagccca cattgcagcc tttgggggac atcctcagtg cctgggtctgg ctgattcaag     480
caggagccaa cattaacaaa ccggttgtg aggggtgaaac tccattcac aaggcagctc     540
gctctgggag cctagaatgc atcagtgcct ttgtggcgaa tggggctcac gtcgacctga     600

```

149

gaaatgccag	tggcctgaca	gcagcagaca	ttgcacaaac	ccagggtttc	caagagtgtg	660
cccagtttct	cttgaacctc	cagaattgtc	atctgaacca	tttctataac	aatggcatct	720
taaatggggg	tcatcagaat	gtatttccta	atcatattag	tgtgggaaca	aatcgaaaga	780
gatgcttggg	agactcagaa	gactttggag	taaagaaagc	tagaactgaa	gctcaaagct	840
tggattctgc	cgtgccactc	acgaatggcg	acacagaaga	cgatgctgac	aaaatgcacg	900
ttgataggga	gtttgctggt	gtaacagggtg	ggagtggaca	gtttcctggt	agctgcaaca	960
acaatccaat	ggttgaagac	accaaacagc	aggagagtgg	ttctgttgga	ccaaaagaaa	1020
tagaaatata	tactgtgtca	gcaatgcaga	ccccctgtcg	ttgcaggaat	caatatgaaa	1080
aacagtagct	ccgtatcgaa	tacattgaca	aatggatgtg	tcatcaatgg	acatttggac	1140
ttccccctca	cgacccccgt	cagtgggatg	gaaagcagga	atggccagtg	cttgacagga	1200
actaacggaa	ttagcagtgg	attagcccca	ggacagccgt	ttccgagtag	ccagggttct	1260
ctctgcatta	gtgggactga	ggagccagag	aagaccctga	gagctaacc	tgagttgtgc	1320
ggttctctgc	acctgaacgg	gagtccaagt	agctgcatag	ccagtaggcc	ttcctgggtg	1380
gaagacattg	gggataacct	gtactatgga	cactaccacg	ggtttgggga	cactgctgaa	1440
agcatcccg	aactgaacag	tgtggtcgag	cattccaagt	ccgtgaaggt	gcaggagcgg	1500
tacgacagtg	ccgtgctggg	caccatgcac	ctgcaccacg	gctcctagag	acgtgacct	1560
ggctctcgga	aacgcaggag	tccttcctgg	tagccagctc	agaataccca	tgtagcagca	1620
acttgaacga	atgtcacaac	ttgtactttt	tttatatact	tcaactttct	gaaaaagtaa	1680
acttcgacaa	gttcccagca	actgcttggt	tgtgcatgag	tagggcttac	taagtgcata	1740
gatgtttcta	cagtgagggtg	tcctttttat	aagggtgact	tttggagttt	ttctgatgcc	1800
aatctcaaca	ttgtcttttt	aatactgtca	ccagatattg	ccatttttct	ttttgttaaa	1860
agattatatg	atcaagataa	attgggggtg	taaatcaggt	gcctggtaat	ttatctcttt	1920
gcacatgggc	atcattttta	aaagcttgct	tccactcttt	tctgtagaat	ttgacggaac	1980
acagctattt	ccctatgcaa	ggtacagcct	tacaaagatt	tctgcagtga	tttgtgtgaa	2040
gaagagaatg	tttgtctttt	tcaatgaagc	tttgcagatc	accatgtggg	tgaagggttt	2100
agttgtggac	acagtgggtcc	ctccttaatg	atgaagatca	ctgccttggg	cttcatggaa	2160
aacaggccca	gcctgggggt	gcgtttggat	ttattgtttt	tattccacac	ttcctacttg	2220
gtctctggaa	gttttaccac	atgtaacaga	ttcctttata	tgtagtga	atcactat	2280
gtagaaactg	tcagggtcaaa	atatttaact	gactgttgac	atgtattttc	ttttttcctt	2340
gtttttgttt	tttgggggtt	tctgctttta	gatatatacc	actatgtata	tccagttaac	2400
tgagagaatt	ttgactctct	taataaaaact	gcattaagtt	tttgattttg	tagaaaattag	2460

150

cttttgtcta ggcaactagt ggttatactc tgcaaatatt gtaatgaatt tttacttttt 2520
 tgatttttgt aataaaaatt ggtgcagata aaatgtcaaa tgaacaaacc agtgttctaa 2580
 gagtgttact aacattttgt tctaaaactg tccttcacaa attgaataaa aaactctcac 2640
 actca 2645

<210> 106
 <211> 2439
 <212> DNA
 <213> Homo sapien

<400> 106
 gcgcgcgcgc gttagggtaa tccatttttc tgtttgtttc tgaattgctt tggagaaggg 60
 aactcagact ccaaggaata caacagctgc cgtgtgaagc ttttctcact tttccagtgg 120
 tcgtagcagg aacagcgtct tgacctcaca tcagtgcagt gaaaagagct gagtttgaag 180
 tgcataattgg tcttcataa agaatgtgtg tatgtgctgt ttttttaatt aagtctggta 240
 atttatttga tacatgtatt gtagttttag ttgaattaaa ctagttagta aaactgttgg 300
 tgaggtagtg tttcagttac tctgtagata agctcttttg ttaaacctc ttaagttttt 360
 tgttcctgaa tatgttaca acatctatct tctcttactg tagctatcac atgaggaaaa 420
 aaacttgaaa aaaagacttc atagtgtgtg tttatgtttg attttagaag acccttttga 480
 ggacttcttt gggaaatcgaa ggggtccccg aggaagcaga agccgagggg cggggtcgtt 540
 tttctctgcg ttcagtggat ttccgtcttt tggaagtggg ttttcttctt ttgatacagg 600
 atttacttca tttgggtcac taggtcacgg gggcctcact tcattctctt ccacgtcatt 660
 tggtggtagt ggcatgggca acttcaaact gatatcaact tcaactaaaa tggttaatgg 720
 cagaaaaatc actacaaaga gaattgtcga gaacggtcaa gaaagagtag aagttgaaga 780
 agatggccag ttaaagtcct taacaataaa tgggtgtggc gacgacgatg ccctcgctga 840
 ggagcgcgat cggagaggcc agaacgccct gccagcccag cctgccggcc tccgcccgcc 900
 gaagccgccc cggcctgcct cgctgctgag acacgcgcct cactgtctct ctgaggagga 960
 gggcgagcag gaccgacctc gggcaccccg gccctgggac cccctcgctt ccgcagcagg 1020
 attgaaagaa ggtggcaaga ggaagaagca gaagcagaga gaggagtcga agaagaagaa 1080
 gtcgacaaaa ggcaatcact agaccggact tgaggcacgc ggtgcacccc cagacgctgg 1140
 cgctccaccg tgctcggcat gcggtcgtgc acacgcgcta ggtagcagcg tcggtcagga 1200
 ctgtctcgag gccacactcg ctccggcagga ttatgcgatc acggatcagt cagagcaggg 1260
 tcaggagacg gggctgacgg cacgggtggc ggggacagac gtttgggact tggccgcgac 1320
 tctctgcttc tctccagctc tcaatctgct gcattttcct ctagtgcttc cggatcctct 1380

151

```

tcattctttt cggctactca accactccgc atgctgctgg aatatttctg gctttagaag 1440
tacaggaggg cgcagatggc taactgagta acattcatga aatgaggctt tctgtggcgg 1500
cgtagtgttt ggaattagaa ggtaattcag tagagtgtaa cttagagaat attgcaagtg 1560
acacattgaa tcctgcccgt cagggcacct tttcctcaga gcaatccggc cacacgaata 1620
gaaggctgtc gtgaatcaca tcagatgtaa aatcattcct tctgtttact cttttaattt 1680
tcaccccttg caggtagtgc aaattcaact tcaaatatgg tgtaggtttt gctagattcc 1740
atattttttt cttggatttt tgctaattat ttttagcaaa aaatttttgc tcagtggcac 1800
cctccctagt gtccatgggt tagggccatg ctggggaaaa cgggccggta tttacacacg 1860
cgcaaaacac ccagagacgg cacaaggagg ttgaactcat gtttcagttc gcgaacattg 1920
actccttacg aaagtcactt cattctaact agatgcgcc acttccggtc attatttctg 1980
ttgcatgatg tattgcttct tcacgttttg tttttattga gcacggagta gaattccagg 2040
gctgccttga cttcttccct gcatgctccc tcccagtgac tttccttccc tttcacatga 2100
ggatctgccg ttcattgttg tttctccttt gtctcttgg acttgagggc attgtgaaaa 2160
gctttgctgt gatttaaaaa tgccagcaat ttaaatctag cagtgttgaa gctgggaatt 2220
ttttggcgca atccatgtag cagtgaccca ggcttgggag ccagaaacaa gtgtgacctg 2280
ggattttatt taacacaact gttgccaaag agttggcttt gtttatttgg ttttggcggg 2340
gagaggagtg gtatttgatg ctttctgtgg acaatgtaac cctaaacaca tcatgtattt 2400
taaatgccac ctacataaat aaaacataag catattgaa 2439

```

```

<210> 107
<211> 1958
<212> DNA
<213> Homo sapien

```

```

<400> 107
ggcagggcct cagggcccca agagggctgt gagcaggcac cttcctgtga ggcacgatgc 60
caccatgtga agttgtgtgg gtctttggag agttctggca ggttgagggc ctgtgcgtgt 120
gggaccaggg cctgccctgc ctgctgccgg ccatgttccc tctgccaggt catgttgaca 180
attgtggcca aattcaacat cgggcctgtt tcctgcaccc acgttggcac aactgagtt 240
caagacagtt tctgcagaaa tgaatccaga ggtgtttatt tctggcacgt tggatcagac 300
gttcagaggg tgggcgtctt tctccatgca cttcagctgg atcagtcagt gtcagccatt 360
gacttttagca acagggaaat gaaagacttt cctcaataag aaatgggact ttctcacgct 420
ggatggcaca gagctggcaa acatgaggcc agtccacacc aaggctttgc ctgcagaaag 480
gctgcactgt ggccagcggg ggaagcggag gagacacctg tggacaccct cccacagga 540

```

152

```

ttgaaagaag gtggcaagag gaagaagcag aagcagagag aggagtcgaa gaagaagaag      600
tcgaccaaag gcaatcacta gaccggactt gaggcacgcg gtgcaccccc agacgctggc      660
gctccaccgt gctcggcatg cggtcgtgca cacgcgctag gtagcagcgt cggtcaggac      720
tgtctcgagg ccacactcgc tcggcaggat tatgcatca cggatcagtc agagcagggc      780
caggagacgg ggctgacggc acgggtggcg gggacagacg tttgggactt ggccgcgact      840
ctctgcttct ctccagctct caatctgctg cattttcctc tagtgcttcc ggatcctctt      900
cattcttttc ggctactcaa ccactccgca tgctgctgga atatttctgg ctttagaagt      960
acaggagggc gcagatggct aactgagtaa cattcatgaa atgaggcttt ctgtggcggc     1020
gtagtgtttg gaattagaag gtaattcagt agagtgtaac ttagagaata ttgcaagtga     1080
cacattgaat cctgcccgtc agggcacctt ttctcagag caatccggcc acacgaatag     1140
aaggctgtcg tgaatcacat cagatgtaaa atcattcctt ctgtttactc ttttaatttt     1200
catcctttgc aggtagtgc aattcaactt caaatatggg gtaggttttg ctagattcca     1260
tatttttttc ttggattttt gctaattatt tttagcaaaa aatttttgct cagtggcacc     1320
ctccctagtg tccatgggtt agggccatgc tggggaaaac gggccgggtat ttacacacgc     1380
gcaaaacacc cagagacggc acaaggaggt tgaactcatg tttcagttcg cgaacattga     1440
ctccttacga aagtcacttc attctaacta gatgcgcccc cttccgggtc ttatttcggt     1500
tgcatgatgt attgcttctt cacgttttgt ttttattgag cacggagtag aattccaggg     1560
ctgccttgac ttcttccctg catgctccct cccagtgaact ttcttccct ttcacatgag     1620
gatctgccgt tcatgttgct ttctcctttg tctctttgga cttgagggca ttgtgaaaag     1680
ctttgctgtg atttaaaaat gccagcaatt ttaatctagc agtggtgaag ctgggaattt     1740
tttggcgcaa tccatgtagc agtgaccag gcttgggagc cagaaacaag tgtgacctgg     1800
gattttattt aacacaactg ttgccaaaga gttggctttg tttatttggt tttggcgggg     1860
agaggagtgg tatttgatgc tttctgtgga caatgtaacc ctaaacacat catgtatttt     1920
aaatgccacc tacataaata aaacataagc atattgaa                                1958

```

```

<210> 108
<211> 3400
<212> DNA
<213> Homo sapien

```

```

<400> 108
ccgccccgtc gcgcagcgt gcggagaaac ctggcgccgc gaccagtgtg tcgggggtct      60
tggaatggc taagaaagct ggcctctgcc taggaggctc tcggcagggg ggctgccaat     120
ccgggatggg taccggaaat gaaccgcgcg atctggccct ttcacacca cttagctttg     180

```

taggggggttt gtagtcttta caggacaaaa gggcacctat tccagagcac accccctttt	240
cttcctctcc cttctgtgcc tcaactgcttt ccgatctcat agttgctccc agtaagtgct	300
ccaaagtcgt gtggggccgc atctgaggtc tgcccgcctg aattgggttg gaggggttca	360
tgttctgaga aggcgaaggc caggggtgtag gatggtgccc cgtgactttc cagtctttgt	420
aggcgtgcct ggtatcttca tcgtaatagc agagtcgtca actttctatg aaagggtggt	480
ttaccttgat aggtggtagt ggtggcctaa tttgtatcat tgtcatgggt ttagtttcca	540
tattctagtt agcagctgct actgttctct tgcaagtgtg taatgtgaaa taataaacgt	600
gaagacttct gaataagggt attctgtgtc tttcatagta gaaaccttaa tgatcgggtct	660
gttgtagtga aactctttta aaaggcgcta tagaaaacca atttctgagt aaaccagcag	720
acagcatgac ttgtaaatgg tcttttaatt aattaaaaag aaattagtca gctacaagca	780
tgaacatgtg gaacgcttac ctttgtacta ggcgtttttg tttttgtttt aatggctttt	840
ggaatattat agtattaaca tctggaaaac taggtaaatt tatcttagaa ttaagttttt	900
tgctcctttt ttgcagaaaa agaacagcaa gaagcgattg aacacattga tgaagtacaa	960
aatgaaatag acagacttaa tgaacaagcc agtgaggaga ttttgaaagt agaacagaaa	1020
tataacaaac tccgccaacc attttttcag aagaggctcag aattgatcgc caaaatccca	1080
aatttttggg taacaacatt tgtcaaccat ccacaagtgt ctgcactgct tggggaggaa	1140
gatgaagagg cactgcatta tttgaccaga gttgaagtga cagaatttga agatattaaa	1200
tcagggttaca gaatagattt ttattttgat gaaaatcctt actttgaaaa taaagttctc	1260
tccaaagaat ttcacttgaa tgagagtggg gatccatctt cgaagtccac cgaaatcaaa	1320
tggaaatctg gaaaggattt gacgaaacgt tcgagtcaaa cgcagaataa agccagcagg	1380
aagaggcagc atgaggaacc agagagcttc tttacctggt ttactgacca ttctgatgca	1440
ggtgctgatg agttaggaga ggtcatcaaa gatgatattt ggccaaaccc attacagtac	1500
tacttggttc ccgatatgga tgatgaagaa ggagaaggag aagaagatga tgatgatgat	1560
gaagaggagg aaggattaga agatattgac gaagaagggg atgaggatga aggtgaagaa	1620
gatgaagatg atgatgaagg ggaggaagga gaggaggatg aaggagaaga tgactaaata	1680
gaacactgat ggattccaac cttccttttt ttaaattttc tccagtccct gggagcaagt	1740
tgcagtcttt tttttttttt tttttttttt ccctcttggt ctcagtcgcc ctgttcttga	1800
ggtctctttt ctctacteca tggttetcaa tttatttggg gggaaatacc ttgagcagaa	1860
tacaatggga aaagagtctc tacccttttc tgttcgaagt tcatttttat cccttcctgt	1920
ctgaacaaaa actgtatgga atcaacacca ccgagctctg tgggaaaaaa gaaaaacctg	1980

154

```

ctcccttcgc tctgctggaa gctggaggggt gctaggcccc tgtgtagtag tgcatagaat 2040
tctagctttt ttctctcttt ctctgtatat tgggctcaga gagtacactg tgtctctatg 2100
tgaatatgga cagttagcat ttaccaacat gtatctgtct actttctctt gtttaaaaaa 2160
agaaaaaaaa acttaaaaaa atgggggttat agaaggtcag caaagggtgg gtttgagatg 2220
tttgggtggg ttaagtgggc attttgacaa catggcttct cctttggcat gtttaattgt 2280
gatatttgac agacatcctt gcagtttaag atgacacttt taaaataaat tctctcctaa 2340
tgatgacttg agccctgccca ctcaatggga gaatcagcag aacctgtagg atcttatttg 2400
gaattgacat tctctattgt aattttgttc ctgtttatct ttaaattttc tttttgtttc 2460
actggaaagg aaagatgatg ctgagtttta aacgttaaaa gtgtacaagt tgctttgtta 2520
caataaaaact aaatgtgtac acaaaggatt tgatgctttt ctctcagcat aggtatgctt 2580
actatgacct tccaagtttg acttgataaa catcactgtc aaactttgtc accctaactt 2640
cgtatttttt gatacgactt ttgcaggatg acctcagggc tatgtggatt gagtaatggg 2700
atttgaatca atgtattaat atctccatag ctgggaaacg tgggttcaat ttgccattgg 2760
tttctgaaag tattcacatc atttgggata ccagatagct caatactctc tgagtacatt 2820
gtgcccttga tttttatctc caagtggcag tttttaaaat tggcctttta cctggatata 2880
aattaattgt gcctgccacc accatccaac agacctgggt ctotaatgcc aagttataca 2940
cgggacagtt gctggcatgt cttcattggc tatataaaat gtggccaaga agataggctc 3000
tcagtaagaa gtctgatggg gagcagtaac tgtccctgct ttctgggtata aagctctcaa 3060
atgtgaccat gtgaatctgg gtgggataat ggactcagct ctgtctgctc aatgccattg 3120
tgagagaag caccctaatt cataagcttt ttaatgctgt aaaatatagt cgctgaaatt 3180
aaatgccact ttttcagagg tgaattaatg gacagtctgg tgaacttcaa aagctttttg 3240
atgtataaaa cttgataaat ggaactattc catcaatagg caaaagtgt acaacctatc 3300
tagatggata gtatgtaatt tctgcacagg tctctgttta gtaaatacat cactgtatac 3360
cgatcaggaa tcttgctcca ataaaggaac ataaagattt 3400

```

<210> 109
<211> 2775
<212> DNA
<213> Homo sapien

```

<400> 109
ccgccccgtc gcgccagcgt gcggagaaac ctggcgccgc gaccagtgt tgggggtct 60
tggaaatggc taagaaagct ggctctgcc taggaggctc tcggcagggg ggctgccaat 120
ccgggatggg taccggaaat gaaccgcgcg atctggccct ttcacacca cttagctttg 180

```

155

taggggggttt gtagtcttta caggacaaaa gggcacctat tccagagcac accccctttt	240
cttcctctcc cttctgtgcc tcaactgcttt ccgatctcat agttgctccc aaaaaagaac	300
agcaagaagc gattgaacac attgatgaag tacaaaatga aatagacaga cttaatgaac	360
aagccagtga ggagattttg aaagtagaac agaaatataa caaactccgc caaccatttt	420
ttcagaagag gtcagaattg atcgccaaaa tcccaaattt ttgggtaaca acatttgtca	480
accatccaca agtgtctgca ctgcttgggg aggaagatga agaggcactg cattatttga	540
ccagagttga agtgacagaa tttgaagata ttaaatacagg ttacagaata gattttttatt	600
ttgatgaaaa tccttacttt gaaaataaag ttctctccaa agaatttcat ctgaatgaga	660
gtgggtgatcc atcttcgaag tccaccgaaa tcaaataaggaa atctggaaag gatttgacga	720
aacgttcgag tcaaacgcag aataaagcca gcaggaagag gcagcatgag gaaccagaga	780
gcttcttttac ctgggtttact gaccattctg atgcaggtgc tgatgagtta ggagagggtca	840
tcaaagatga tatttggcca aaccatttac agtactactt gggtcccgat atggatgatg	900
aagaaggaga aggagaagaa gatgatgatg atgatgaaga ggaggaagga ttagaagata	960
ttgacgaaga aggggatgag gatgaagggtg aagaagatga agatgatgat gaaggggagg	1020
aaggagagga ggatgaagga gaagatgact aaatagaaca ctgatggatt ccaaccttcc	1080
ttttttttaa ttttctccag tccctgggag caagttgcag tctttttttt tttttttttt	1140
tttttccctc ttgtgctcag tcgccctgtt cttagggtct cttttctcta ctccatgggt	1200
ctcaatttat ttggggggaa ataccttgag cagaatacaa tgggaaaaga gtctctaccc	1260
ctttctgttc gaagttcatt tttatccctt cctgtctgaa caaaaactgt atggaatcaa	1320
caccaccgag ctctgtggga aaaaagaaaa acctgtcccc ttctgtctgc tggaagctgg	1380
agggtgctag gccctgtgt agtagtgcag agaattctag cttttttcct cctttctctg	1440
tatatgggc tcagagagta cactgtgtct ctatgtgaat atggacagtt agcatttacc	1500
aacatgtatc tgtctacttt ctcttgttta aaaaaagaaa aaaaaactta aaaaaatggg	1560
gttatagaag gtcagcaaag ggtgggtttg agatgtttgg gtgggttaag tgggcatttt	1620
gacaacatgg cttctccttt ggcatgttta attgtgatat ttgacagaca tccttgcagt	1680
ttaagatgac acttttaaaa taaattctct cctaataatg acttgagccc tgccactcaa	1740
tgggagaatc agcagaacct gtaggatctt atttggaatt gacattctct attgtaattt	1800
tgttcctgtt tattttttaa ttttcttttt gtttactgg aaaggaaaga tgatgctcag	1860
ttttaaacgt taaaagtgt caagttgctt tgttacaata aaactaaatg tgtacacaaa	1920
ggatttgatg cttttctctc agcataggta tgcttactat gaccttccaa gtttgacttg	1980
tataacatca ctgtcaaact ttgtcaccct aacttcgtat tttttgatac gcactttgca	2040

156

ggatgacctc agggctatgt ggattgagta atgggatttg aatcaatgta ttaatatctc 2100
 catagctggg aaacgtgggt tcaatttgcc attggtttct gaaagtattc acatcatttg 2160
 ggataccaga tagctcaata ctctctgagt acattgtgcc cttgattttt atctccaagt 2220
 ggagttttt aaaattggcc ttttacctgg atataaatta attgtgcctg ccaccaccat 2280
 ccaacagacc tgggtgctcta atgccaagtt atacacggga cagttgctgg catgtcttca 2340
 ttggctatat aaaatgtggc caagaagata ggctctcagt aagaagtctg atggtgagca 2400
 gtaactgtcc ctgctttctg gtataaagct ctcaaagtgt accatgtgaa tctgggtggg 2460
 ataatggact cagctctgtc tgctcaatgc cattgtgcag agaagcacc ctaatgcataa 2520
 gctttttaat gctgtaaaat atagtcgctg aaattaaatg ccactttttc agaggtgaat 2580
 taatggacag tctggtgaac ttcaaaagct ttttgatgta taaaacttga taaatggaac 2640
 tattccatca ataggcaaaa gtgtaacaac ctatctagat ggatagtatg taatttctgc 2700
 acaggtctct gtttagtaaa tacatcactg tataccgatc aggaatcttg ctccaataaa 2760
 ggaacataaa gattt 2775

<210> 110
 <211> 4476
 <212> DNA
 <213> Homo sapien

<400> 110
 catcttgtag ccggcgggcg ggattgaatg agcccgccga gcccgggccg ccgtcgggag 60
 cagcgagggc cgcgagccgc cgccaccatg gccacgccgc tgcccgccg ccggggcggg 120
 cccgccacgc cgctgtcgcc cacgcgcctg tcgcggtgc aggagaagga ggagctgcgc 180
 gagctcaacg accgcctggc gcactacatc gaccgcgtcc gcgcgctgga gctggagaac 240
 gaccggctcc tgctcaagat ctgagagaag gaggaggtga ccacgcgcga ggtgagtggt 300
 atcaaggcgc tgtacgagtc ggagctggcc gatgcccgga ggtcctgga tgagacggct 360
 cgagagcgtg cccggctgca gatagagatt gggaagctga gggcagagtt ggacgaggtc 420
 aacaagagcg ccaagaagag ggagggcgag cttacgggtg cccagggccg tgtgaaggac 480
 ctggagtcct tgttccaccg gagcgaggtg gagctggcag ctgccctcag cgacaagcgc 540
 ggcttgagga gtgacgtggc tgagctgcgg gccagctgg ccaaggccga ggacgggtcat 600
 gcagtggtcca aaaagcagct ggagaaggag acgctgatgc gtgtggacct ggagaaccgc 660
 tgccagagcc tgcaggagga gctggacttc cggaagagtg tgttcgagga ggaggtgcgg 720
 gagacgcggc ggcggcacga gcggcgctg gtggaggtgg acagcagccg gcagcaggag 780
 tacgacttca agatggcaca ggcgctggag gagctgcgga gccagcacga cgagcaagtg 840

157

cggctctaca agctggagct ggagcagacc taccaggcca agctggacag cgccaagctg	900
agctctgacc agaacgacaa ggcggccagt gcggtctgcg aggagctgaa ggaggcccg	960
atgcgcctgg agtccctcag ctaccagctc tccggcctcc agaagcaggc cagtgccgct	1020
gaagatcgca ttcgggagct ggaggaggcc atggccgggg agcgggacaa gttccggaag	1080
atgctggacg ccaaggagca ggagatgacg gagatgcggg acgtgatgca gcagcagctg	1140
gccgagtacc aggagctgct ggacgtgaag ctggccctgg acatggagat caacgcctac	1200
cggaagctcc tggaggggcg ggaggagagg ctgaagctgt cccccagccc atcctcgcg	1260
gtcaccgtct cacgagccac ctcgagcagc agcggcagct tgtccgccac cgggcgcctg	1320
ggccgcagta agcgggaagcg gctggagggtg gaggagccct tgggcagcgg cccaagcgtc	1380
ctgggcacgg gcacgggtgg cagcgggtgg ttccacctgg ccagcaggc ctgggcctcg	1440
ggtagcgtca gcatcgagga gatcgacctg gagggcaagt ttgtgcagct caagaacaac	1500
tccgacaagg atcagtctct ggggaactgg agaatacaaga ggcaggtctt ggagggggag	1560
gagatcgct acaagttcac gcccaagtac atcctgcgcg ccggccagat ggtcacgggtg	1620
tgggcagctg gtgcgggggt ggcccacagc cccccctcga cgctggtgtg gaagggccag	1680
agcagctggg gcacggggcg gagcttccgc accgtcctgg ttaacgcgga tggcgaggaa	1740
gtggccatga ggactgtgaa gaagtctcg gtgatgcgtg agaatagaa tggggaggaa	1800
gaggaggagg aagccgagtt tggcgaggag gatcttttcc accaacaggg ggaccgagg	1860
accacctcaa gaggtgcta cgtgatgtga acccacactc ctcatccaca cacctttctt	1920
taccagagc cactgaaaac tatttttata tcattggctt tctttagtgc ttgatacatt	1980
tctagagaat ttctaagcg actgccagaa cgtgtgggtg ggtctcccc agccctccct	2040
cctggcgggt ctccctcagc ctcaactcgc tgccacttcg ccgctgcccc ggagactttt	2100
caatcccacc ccactcctca tctcaccatt tggtaaatt ggaagcccag ggccaggacc	2160
cggagggtta gaagatgctt gggcttgagg ggaggaggcg cggcgaggct agcaggggga	2220
caggagacgg ccctgctgcg gacggagcgc ggaaactgcg taggaattca gtggtggtgg	2280
gtttttttaa ggctttctac aaaaccaa atcagaatcca ggcgtcgacc tgggtggggc	2340
cggggccaag cctgcattct ggctgccag ctccggacag cgggaactcc tcaggcagcc	2400
acgcagcggg tgtgggccag catggggatg gcgtggcccc cagggcgggt tttcactccg	2460
ctgcctgggc ttccagattc cgttctggc agcgcaccgg ccgggtttct cggaccgttg	2520
actttatttg ggggagtttt ccgcagttc agttcctgac tgtgcaaggc caacagggca	2580
ggggagggga agacctgggg aaggaagaat gaggacagtc ccgtcgtaag acctgtcaca	2640

158

acaataagca	gggaggggag	atgtggaggg	gacacatctg	gttgcccttg	aggcagaagc	2700
tgtgagtttc	agaacagctg	tctgcagggg	acgccaccat	gttgaccctc	tggaggagag	2760
cgctgtggag	cccctcccgt	gttccagctc	cgtctgccct	gtgcctatat	atacacatgc	2820
gtctatccat	actgtgcttt	tatctgtgat	tttctcgtg	aaaccatgtt	tctcagacag	2880
gccaaaggcca	cctgactcct	atcacgacgc	acccaagccc	ctcagtccag	cttcccaatg	2940
cctggcaccc	cccttcggca	atagctcacc	gtttacaccc	tccctcatag	atacacagaa	3000
gttatttttt	taatggatat	ttattttttt	acattgggtca	gtacacaggt	caggagctca	3060
cgccagggcc	ttgaggacag	gctgaccctc	ctccccgggg	tggcgtgggg	ctggggcacc	3120
ctccgacggc	agagcctcct	tcagaaagtg	cagctcaagt	cttaaagaca	ccaaaactga	3180
gccatgggca	cgcgccgtct	ccgggccatg	gcgttcaactg	cagggcgggg	gcggcaccgc	3240
tccccatgta	gatagagggg	ggcgcatctt	accaaagcat	ttcctcctgg	aggctacgcc	3300
gctgtgctcc	cagtcaggcg	gctggtaggg	agctttgcct	gccccgggga	taccctctgc	3360
cagccgctgg	aagtgggaat	gctggcgaca	gactgtgtcc	tctttccac	cttcatagca	3420
ggaatcacc	ggacccgact	ggctgggctt	cgtgctagcg	agggttctgg	gggtgggtct	3480
tgggtgatctt	gtcctatggg	gagtctgcag	tgggtctcagc	cacatcctat	gtattttggc	3540
tctggaggag	caaagctgta	tcctggagtt	ggctctgtgat	ttgccgacag	ccttgcaggc	3600
tgggctcagg	gacaaagtcc	ccccaaaac	ccgcaggctc	tcattgtccag	acgtgcccc	3660
gtcctgtcct	gaaaacagca	cgccccaggc	ccacagaacc	ccccacccta	catttgcctt	3720
gggtggagct	gggggtggtc	ctaggactgc	gggtgccctt	agctgaagg	ggtggggaga	3780
agcgtggact	gggcagcctg	tgggtaattg	gagggtcatt	gagaattgag	tctttggaaa	3840
cactaagaaa	atcaaatttt	taaaagtatt	ttatggcctg	ggaaacaatt	tgcatattgtc	3900
cccaaatacg	cttagctgtg	tgccgcttag	aacgatgcga	aaccatccct	ctgtgtaagc	3960
ccgtgccgtg	tgactcgaag	cctagcggcc	tccttgcgaa	gcatcagacg	ccaccagcc	4020
ctgggggggag	gccacgcct	gctggaccaa	cgcggttct	ggggtgcaca	gcgccagggt	4080
aacgtgaag	cctgccccgc	tgagcccagg	agccgggagg	cctgcgggct	gaccagaat	4140
ccgatcatgc	acctgtcctc	atgccagcgg	ctttggctgg	ggttggtctg	aagcctgcac	4200
gcggcagttc	tttgttaaag	atctgagggg	ctcctcagtc	ctggggcgtc	gccgcctgca	4260
gcctcttcca	agccctgcgt	ccagcgagcg	tcacagcaca	acctgcaaaa	acggagctgg	4320
gctgcagctg	gggtggcat	ggactttcat	ttcagagatt	cggtttttaa	gaagatgcat	4380
gcctaagtgtg	ttcttttttt	tttccaatga	tttgtaatat	acattttatg	actggaaact	4440
ttttgtaca	acactccaat	aaacattttg	atttta			4476

<210> 111
<211> 1887
<212> DNA
<213> Homo sapien

<400> 111
gcctggtggc gaattcggca cgaggatgag cccgccgagc ccgggccgcc gtcgggagca 60
gcgcaggccg cgagccgccg ccaccatggc cagcccgctg cccggcccg cgggcgggccc 120
cgccacgccg ctgtcgccca cgcgcctgtc gcggtgcag gagaaggagg agctgcgcga 180
gctcaacgac cgctggcgc actacatcga ccgcgtccgc gcgctggagc tggagaacga 240
ccggctcctg ctcaagatct cagagaagga ggaggtgacc acgcgcgagg tgagtggcat 300
caaggcgctg tacgagtcgg agctggccga tgcccggaga gtcttgatg agacggctcg 360
agagcgtgcc cggctgcaga tagagattgg gaagctgagg gcagagttgg acgaggtcaa 420
caagagcgcc aagaagaggg agggcgagct tacggtggcc cagggccgtg tgaaggacct 480
ggagtccctg ttccaccgga gcgaggtgga gctggcagct gccctcagcg acaagcgcg 540
cctggagagt gacgtggctg agctgcgggc ccagctggcc aaggccgagg acggtcatgc 600
agtggccaaa aagcagctgg agaagggtg tccttgctct cagaagtccc ggtcccatgt 660
agatagaggg gggcgcatct taccaaagca tttcctcctg gaggctacgc cgctgtgctc 720
ccagtcaggc ggctggtagg gagctttgcc tgccccgggg ataccctctg ccagccgctg 780
gaagtgggaa tgctggcgac agactgtgtc ctctttccca ccttcatagc agaatcacc 840
cggacccgac tggctgggct tcgtgctagc gagggttctg ggggtgggtc ttggtgatct 900
tgtcctatgg ggagtctgca gtggtctcag ccacatccta tgtattttgg ctctggagga 960
gcaaagctgt atcctggagt tggctctgtga tttgccgaca gccttgacagg ctgggctcag 1020
ggacaaagtc cccccaaaa cccgcaggtc ctcatgtcca gacgtgccc agtcctgtcc 1080
tgaaaacagc acgccccagg cccacagaac cccccacct acatttgcct tgggtggagc 1140
tgggggtggt cctaggactg cgggtgccct tagctgaagg ggggtggggag aagcgtggac 1200
tgggcagcct gtgggtaatt ggaggttcat tgagaattga gtctttggaa aactaagaa 1260
aatcaaattt ttaaaagtta tttatggcct gggaaacaat ttgcatttgt ccccaaatac 1320
gcttagctgt gtgccgctta gaacgatgcg aaaccatccc tctgtgtaag ccggtgccgt 1380
gtgactcgaa gcctagcgcc ctccctgcga agcatcagac gccaccagc cctgggggga 1440
ggcccacgcc tgctggacca acgcgggttc tggggtgcac agcgcaggt taacgctgaa 1500
gcctgccccg ctgagcccag gagccgggag gcctgcgggc tgaccagaa tccgatcatg 1560
cacctgtcct catgccagcg gctttggctg ggggtgggtc gaagcctgca cgcggcagtt 1620

160

ctttgttaaa gatctgaggg actcctcagt cctggggcgt cgccgcctgc agcctcttcc 1680
aagccctgcg tccagcgagc gtcacagcac aacctgcaaa aacggagctg ggctgcagct 1740
ggggctggca tggactttca tttcagagat tcggttttta agaagatgca tgcctaattgt 1800
gttctttttt ttttccaatg atttgtaata tacattttat gactggaaac ttttttgtac 1860
aacactccaa taaacatttt gatttta 1887

<210> 112
<211> 4041
<212> DNA
<213> Homo sapien

<400> 112
catcttgacg ccggcggcgc ggattgaatg agcccgcga gcccgggccg ccgtcgggag 60
cagcgcaggc cgcgagccgc cgccaccatg gccacgcgc tgcccgccg cgcgggcggg 120
cccgccacgc cgctgtcgcc cacgcgcctg tcgcggctgc aggagaagga ggagctgcgc 180
gagctcaacg accgcctggc gcactacatc gaccgcgtcc gcgcgctgga gctggagaac 240
gaccggctcc tgctcaagat ctcaagaaag gaggaggtga ccacgcgcga ggtgagtggc 300
atcaaggcgc tgtacgagtc ggagctggcc gatgcccgga gagtccctga tgagacggct 360
cgagagcgtg cccggctgca gatagagatt gggaagctga gggcagagtt ggacgaggtc 420
aacaagagcg ccaagaagag ggagggcgag cttacggtgg cccagggccg tgtgaaggac 480
ctggagtccc tggtccaccg gagcgaggtg gagctggcag ctgccctcag cgacaagcgc 540
ggcctggaga gtgacgtggc tgagctgcgg gccagctgg ccaaggccga ggacggcat 600
gcagtggcca aaaagcagct ggagaaggag acgctgatgc gtgtggacct ggagaaccgc 660
tgccagagcc tgcaggagga gctggacttc cggaagagtg tggtcgagga ggaggtgcgg 720
gagacgcggc ggcggcacga gcggcgctg gtggaggtgg acagcagccg gcagcaggag 780
tacgacttca agatggcaca ggcgctggag gagctgcgga gccagcacga cgagcaagtg 840
cggctctaca agctggagct ggagcagacc taccaggcca agctggacag cgccaagctg 900
agctctgacc agaacgacaa ggcggccagt gcggctcgcg aggagctgaa ggaggcccgc 960
atgcgcctgg agtccctcag ctaccagctc tccggcctcc agaagcaggc cagtgccgct 1020
gaagatcgca ttcgggagct ggaggaggcc atggccgggg agcgggacaa gttccggaag 1080
atgctggacg ccaaggagca ggagatgacg gagatgcggg acgtgatgca gcagcagctg 1140
gccagtagcc aggagctgct ggacgtgaag ctggccctgg acatggagat caacgcctac 1200
cggaagctcc tggagggcga ggaggagagg ctgaagctgt ccccagccc atcctcgcgc 1260
gtcaccgtct cagagccac ctcgagcagc agcggcagct tgtccgccac cgggcgcctg 1320

ggccgcagta agcggaagcg gctggagggtg gaggagccct tgggcagcgg cccaagcgtc	1380
ctggggcacgg gcacgggtgg cagcgggtggc ttccacctgg cccagcaggc ctcggcctcg	1440
ggtagcgtca gcatcgagga gatcgacctg gagggcaagt ttgtgcagct caagaacaac	1500
tccgacaagg atcagtctct ggggaactgg agaatacaaga ggcaggctct ggagggggag	1560
gagatcgccct acaagttcac gccaagtac atcctgcgcg ccggccagat ggtcacgggtg	1620
tgggcagctg gtgcgggggt ggcccacagc cccccctcga cgctggtgtg gaagggccag	1680
agcagctggg gcacgggcga gagcttcgcg accgtcctgg ttaacgcgga tggcgaggaa	1740
gtggccatga ggactgtgaa gaagtcctcg gtgatgcgtg agaatagaa tggggaggaa	1800
gaggaggagg aagccgagtt tggcgaggag gatcttttcc accaacaggg ggacccgagg	1860
accacctcaa gaggtgcta cgtgatgtga acccacactc ctcatccaca cacctttctt	1920
taccagagc cactgaaaac tatttttata tcattggctt tctttagttc ttgatacatt	1980
tctagagaat ttctaagcga actgccagaa cgtgtgggtg ggtctcccc agccctccct	2040
cctggcgggt ctctccagc ctcaacttcg tgccacttcg ccgctgcccc ggagactttt	2100
caatcccacc ccactcctca tctcaccatt tgggtcaaatt ggaagcccag ggccaggacc	2160
cggaggttta gaagatgctt gggcttggag ggaggagggc cggcgaggct agcgagggga	2220
caggagacgg ccctgctgcg gacggagcgc ggaaactgcg taggaattca gtggtggtgg	2280
gttttttttaa ggctttctac aaaaccaa atcagaatcca ggcgtcgacc tgggtggggcc	2340
cggggccaag cctgcattct ggctgcccag cttcggacag cgggaactcc tcaggcagcc	2400
acgcagcggg tgtgggcccag catggggatg gcgtggcccc cagggcgggt tttactccg	2460
ctgcctgggc ttccagattc ccgttctggc agcgcaccgg ccgggtttct cggaccgttg	2520
actttatattg ggggagtttt cccgcagttc agttcctgac tgtgcaaggc caacagggca	2580
ggggagggga agacctggg aaggaagaat gaggacagtc ccgtcgtaag acctgtcaca	2640
acaataagca gggaggggag atgtggaggg gacacatctg gttgccttgg aggcagaagc	2700
tgtgagtttc agaacagctg tctgcaggga acgccacat gttgaccctc tggaggagag	2760
cgctgtggag cccctcccgt gttccagctc cgtctgccct gtgcctatat atacacatgc	2820
gtctatccat actgtgcttt tatctgtgat ttctctgctg aaaccatgtt tctcagacag	2880
gccaaggcca cctgactcct atcacgacgc acccaagccc ctcagtccag cttcccaatg	2940
cctggcacc ccttcggca atagctcacc gtttacaccc tccctcatag atacacagaa	3000
gttatTTTTT taatggatat ttatTTTTTT acattggtca gtacacaggc caggagctca	3060
cgccagggcc ttgaggacag gctgaccctc ctccccgggg tggcgtgggg ctggggcacc	3120

162

ctccgacggc agagcctcct tcagaaagtg cagctcaagt cttaaagaca ccaaaactga 3180
 gccatgggca cgcgccgtct ccgggccatg gcgttcactg cagggcgggg gcggcaccgc 3240
 tcccctgtga ctgcatcccg cctccctggg gacctgcctg tggcaggaag gaatgggggg 3300
 ccccgagccc aggccgggaa ggagccagcg gccgacaaag cagaaacacc cgctgctcca 3360
 cgtagccctt gctcgctgtc cttgctctca gaagtcccg tcccatgtag atagaggggg 3420
 gcgcatctta ccaaagcatt tcctcctgga ggctacgccg ctgtgctccc agtcaggcgg 3480
 ctggtaggga gctttgcctg ccccggggat accctctgcc agccgctgga agtggaatg 3540
 ctggcgacag actgtgtcct ctttccacc ttcatagcag gaatcaccg gacccgactg 3600
 gctgggcttc gtgctagcga gggttctggg ggtgggtctt ggtgatcttg tcctatgggg 3660
 agtctgcagt ggtctcagcc acatcctatg tatcttggt ctggaggagc aaagctgtat 3720
 cctggagttg gtctgtgatt tgccgacagc cttgcaggct gggctcaggg acaaagtccc 3780
 ccccaaaacc cgcaggctct cattcctggg gcgtcgccgc ctgcagcctc ttccaagccc 3840
 tgcgccagc gagcgtcaca gcacaacctg caaaaacgga gctgggctgc agctggggct 3900
 ggcattggact ttcatctcag agattcgggt ttaagaaga tgcattgcta atgtgttctt 3960
 ttttttttcc aatgatttgt aatatacatt ttatgactgg aaactttttt gtacaacact 4020
 ccaataaaca ttttgatttt a 4041

<210> 113
 <211> 165
 <212> PRT
 <213> Homo sapien

<400> 113

Gly Gly Arg Asn Arg Gly Arg Ala Gly Ser Gln Gly Gly Arg Gly Gly
 1 5 10 15

Gly Ala Gln Ala Ala Ala Arg Val Asn Arg Gly Gly Gly Pro Ile Arg
 20 25 30

Asn Arg Pro Ala Ile Ala Arg Gly Ala Ala Gly Gly Gly Gly Arg Asn
 35 40 45

Arg Pro Ala Pro Tyr Ser Arg Pro Lys Gln Leu Pro Asp Lys Trp Gln
 50 55 60

His Asp Leu Phe Asp Ser Gly Phe Gly Gly Gly Ala Gly Val Glu Thr
 65 70 75 80

Gly Gly Lys Leu Leu Val Ser Asn Leu Asp Phe Gly Val Ser Asp Ala

163

85 90 95

Asp Ile Gln Glu Leu Phe Ala Glu Phe Gly Thr Leu Lys Lys Ala Ala
 100 105 110

Val His Tyr Asp Arg Ser Gly Arg Ser Leu Gly Thr Ala Asp Val His
 115 120 125

Phe Glu Arg Lys Ala Asp Ala Leu Lys Ala Met Lys Gln Tyr Asn Gly
 130 135 140

Val Pro Leu Asp Ala Ser Tyr Ile Pro Pro Leu Leu Gln Leu Leu Pro
 145 150 155 160

Glu Asp Ser Leu Leu
 165

<210> 114
 <211> 164
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> X=any amino acid

<400> 114

Gly Ala Gly Xaa Ala Pro Gly Arg Leu Gln Gly Gly Arg Gly Gly Gly
 1 5 10 15

Ala Gln Ala Ala Ala Arg Val Asn Arg Gly Gly Gly Pro Ile Arg Asn
 20 25 30

Arg Pro Ala Ile Ala Arg Gly Ala Ala Gly Gly Gly Gly Arg Asn Arg
 35 40 45

Pro Ala Pro Tyr Ser Arg Pro Lys Gln Leu Pro Asp Lys Trp Gln His
 50 55 60

Asp Leu Phe Asp Ser Gly Phe Gly Gly Gly Ala Gly Val Glu Thr Gly
 65 70 75 80

Gly Lys Leu Leu Val Ser Asn Leu Asp Phe Gly Val Ser Asp Ala Asp
 85 90 95

Ile Gln Glu Leu Phe Ala Glu Phe Gly Thr Leu Lys Lys Ala Ala Val

164

100

105

110

His Tyr Asp Arg Ser Gly Arg Ser Leu Gly Thr Ala Asp Val His Phe
 115 120 125

Glu Arg Lys Ala Asp Ala Leu Lys Ala Met Lys Gln Tyr Asn Gly Val
 130 135 140

Pro Leu Asp Ala Ser Tyr Ile Pro Pro Leu Leu Gln Leu Leu Pro Glu
 145 150 155 160

Asp Ser Leu Leu

<210> 115
 <211> 256
 <212> PRT
 <213> Homo sapien

<400> 115

Met Ala Leu Arg Val Thr Arg Asn Ser Lys Ile Asn Ala Glu Asn Lys
 1 5 10 15

Ala Lys Ile Asn Met Ala Gly Ala Lys Arg Val Pro Thr Ala Pro Ala
 20 25 30

Ala Thr Ser Lys Pro Gly Leu Arg Pro Arg Thr Ala Leu Gly Asp Ile
 35 40 45

Gly Asn Lys Val Ser Glu Gln Leu Gln Ala Lys Met Pro Met Lys Lys
 50 55 60

Glu Ala Lys Pro Ser Ala Thr Gly Lys Val Ile Asp Lys Lys Leu Pro
 65 70 75 80

Lys Pro Leu Glu Lys Val Pro Met Leu Val Pro Val Pro Val Ser Glu
 85 90 95

Pro Val Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Val Lys Glu
 100 105 110

Glu Lys Leu Ser Pro Glu Pro Ile Leu Val Asp Thr Ala Ser Pro Ser
 115 120 125

Pro Met Glu Thr Ser Gly Cys Ala Pro Ala Glu Glu Asp Leu Cys Gln
 130 135 140

165

Ala Phe Ser Asp Val Ile Leu Ala Val Asn Asp Val Asp Ala Glu Asp
 145 150 155 160

Gly Ala Asp Pro Asn Leu Cys Ser Glu Tyr Val Lys Asp Ile Tyr Ala
 165 170 175

Tyr Leu Arg Gln Leu Glu Glu Glu Gln Ala Val Arg Pro Lys Tyr Leu
 180 185 190

Leu Gly Arg Glu Val Thr Gly Asn Met Arg Ala Ile Leu Ile Asp Trp
 195 200 205

Leu Val Gln Val Gln Met Lys Ser Val Cys Ala Gly Pro Val Cys Gly
 210 215 220

Pro Ile Asp Gly Pro Ala Lys Leu Gly Ala Gln Ile Ala Gly Gly Pro
 225 230 235 240

Ala Val Trp Pro Leu Lys Gly Pro Arg Gly Arg Trp Gly Thr Leu Ala
 245 250 255

<210> 116
 <211> 250
 <212> PRT
 <213> Homo sapien

<400> 116

Ala Gly Ser Ser Arg Arg Ala Ala Ala Glu Arg Leu Leu Val Ser Ala
 1 5 10 15

Gly Cys Arg Ser Leu Ala Gly Arg Ala Ser Gly Val Leu Leu Leu Pro
 20 25 30

Ala Glu Leu Leu Pro Gly Glu Glu Glu Ala Met Ala Leu Arg Val Thr
 35 40 45

Arg Asn Ser Lys Ile Asn Ala Glu Asn Lys Ala Lys Ile Asn Met Ala
 50 55 60

Gly Ala Lys Arg Val Pro Thr Ala Pro Ala Ala Thr Ser Lys Pro Gly
 65 70 75 80

Leu Arg Pro Arg Thr Ala Leu Gly Asp Ile Gly Asn Lys Val Ser Glu
 85 90 95

Gln Leu Gln Ala Lys Met Pro Met Lys Lys Glu Ala Lys Pro Ser Ala

166

100 105 110

Thr Gly Lys Val Ile Asp Lys Lys Leu Pro Lys Pro Leu Glu Lys Val
 115 120 125

Pro Met Leu Val Pro Val Pro Val Ser Glu Pro Val Pro Glu Pro Glu
 130 135 140

Pro Glu Pro Glu Pro Glu Pro Val Lys Glu Glu Lys Leu Ser Pro Glu
 145 150 155 160

Pro Ile Leu Val Asp Thr Ala Ser Pro Ser Pro Met Glu Thr Ser Gly
 165 170 175

Cys Ala Pro Ala Glu Glu Asp Leu Cys Gln Ala Phe Ser Asp Val Ile
 180 185 190

Leu Ala Val Asn Asp Val Asp Ala Glu Asp Gly Ala Asp Pro Asn Leu
 195 200 205

Cys Ser Glu Tyr Val Lys Asp Ile Tyr Ala Tyr Leu Arg Gln Leu Glu
 210 215 220

Glu Glu Gln Ala Val Arg Pro Lys Tyr Leu Leu Gly Arg Glu Val Thr
 225 230 235 240

Gly Asn Met Arg Ala Ile Leu Ile Asp Trp
 245 250

<210> 117
 <211> 406
 <212> PRT
 <213> Homo sapien

<400> 117

Met Glu Ala Ala Ala Val Thr Val Thr Arg Ser Ala Thr Arg Arg Arg
 1 5 10 15

Arg Arg Gln Leu Gln Gly Leu Ala Ala Pro Glu Ala Gly Thr Gln Glu
 20 25 30

Glu Gln Glu Asp Gln Glu Pro Arg Pro Arg Arg Arg Arg Pro Gly Arg
 35 40 45

Ser Ile Lys Asp Glu Glu Glu Glu Thr Val Phe Arg Glu Val Val Ser
 50 55 60

167

Phe Ser Pro Asp Pro Leu Pro Val Arg Tyr Tyr Asp Lys Asp Thr Thr
65 70 75 80

Lys Pro Ile Ser Phe Tyr Leu Ser Ser Leu Glu Glu Leu Leu Ala Trp
85 90 95

Lys Pro Arg Leu Glu Asp Gly Phe Asn Val Ala Leu Glu Pro Leu Ala
100 105 110

Cys Arg Gln Pro Pro Leu Ser Ser Gln Arg Pro Arg Thr Leu Leu Cys
115 120 125

His Asp Met Met Gly Gly Tyr Leu Asp Asp Arg Phe Ile Gln Gly Ser
130 135 140

Val Val Gln Thr Pro Tyr Ala Phe Tyr His Trp Gln Cys Ile Asp Val
145 150 155 160

Phe Val Tyr Phe Ser His His Thr Val Thr Ile Pro Pro Val Gly Trp
165 170 175

Thr Asn Thr Ala His Arg His Gly Val Cys Val Leu Gly Thr Phe Ile
180 185 190

Thr Glu Trp Asn Glu Gly Gly Arg Leu Cys Glu Ala Phe Leu Ala Gly
195 200 205

Asp Glu Arg Ser Tyr Gln Ala Val Ala Asp Arg Leu Val Gln Ile Thr
210 215 220

Gln Phe Phe Arg Phe Asp Gly Trp Leu Ile Asn Ile Glu Asn Ser Leu
225 230 235 240

Ser Leu Ala Ala Val Gly Asn Met Pro Pro Phe Leu Arg Tyr Leu Thr
245 250 255

Thr Gln Leu His Arg Gln Val Pro Gly Gly Leu Val Leu Trp Tyr Asp
260 265 270

Ser Val Val Gln Ser Gly Gln Leu Lys Trp Gln Asp Glu Leu Asn Gln
275 280 285

His Asn Arg Val Phe Phe Asp Ser Cys Asp Gly Phe Phe Thr Asn Tyr
290 295 300

168

Asn Trp Arg Glu Glu His Leu Glu Arg Met Leu Gly Gln Ala Gly Glu
 305 310 315 320

Arg Arg Ala Asp Val Tyr Val Gly Val Asp Val Phe Ala Arg Gly Asn
 325 330 335

Val Val Gly Gly Arg Phe Asp Thr Asp Lys Val Gly Gly Gly Phe Arg
 340 345 350

Pro Arg Ala Ser Gly Pro Val Pro Pro Leu Gly Pro His Phe Leu Met
 355 360 365

Asp Leu Pro Phe Pro Ser Ala Pro Gln Arg Asn Asp Ser Ser Cys Ser
 370 375 380

Ser Gln Ser Gly Asp Pro Val Ala Leu Arg Asn Arg Cys Pro Ala Pro
 385 390 395 400

Ala Lys Leu Cys Pro His
 405

<210> 118
 <211> 525
 <212> PRT
 <213> Homo sapien

<400> 118

Met Glu Ala Ala Ala Val Thr Val Thr Arg Ser Ala Thr Arg Arg Arg
 1 5 10 15

Arg Arg Gln Leu Gln Gly Leu Ala Ala Pro Glu Ala Gly Thr Gln Glu
 20 25 30

Glu Gln Glu Asp Gln Glu Pro Arg Pro Arg Arg Arg Arg Pro Gly Arg
 35 40 45

Ser Ile Lys Asp Glu Glu Glu Glu Thr Val Phe Arg Glu Val Val Ser
 50 55 60

Phe Ser Pro Asp Pro Leu Pro Val Arg Tyr Tyr Asp Lys Asp Thr Thr
 65 70 75 80

Lys Pro Ile Ser Phe Tyr Leu Ser Ser Leu Glu Glu Leu Leu Ala Trp
 85 90 95

Lys Pro Arg Leu Glu Asp Gly Phe Asn Val Ala Leu Glu Pro Leu Ala
 100 105 110

169

Cys Arg Gln Pro Pro Leu Ser Ser Gln Arg Pro Arg Thr Leu Leu Cys
 115 120 125

His Asp Met Met Gly Gly Tyr Leu Asp Asp Arg Phe Ile Gln Gly Ser
 130 135 140

Val Val Gln Thr Pro Tyr Ala Phe Tyr His Trp Gln Cys Ile Asp Val
 145 150 155 160

Phe Val Tyr Phe Ser His His Thr Val Thr Ile Pro Pro Val Gly Trp
 165 170 175

Thr Asn Thr Ala His Arg His Gly Val Cys Val Leu Gly Thr Phe Ile
 180 185 190

Thr Glu Trp Asn Glu Gly Gly Arg Leu Cys Glu Ala Phe Leu Ala Gly
 195 200 205

Asp Glu Arg Ser Tyr Gln Ala Val Ala Asp Arg Leu Val Gln Ile Thr
 210 215 220

Gln Phe Phe Arg Phe Asp Gly Trp Leu Ile Asn Ile Glu Asn Ser Leu
 225 230 235 240

Ser Leu Ala Ala Val Gly Asn Met Pro Pro Phe Leu Arg Tyr Leu Thr
 245 250 255

Thr Gln Leu His Arg Gln Val Pro Gly Gly Leu Val Leu Trp Tyr Asp
 260 265 270

Ser Val Val Gln Ser Gly Gln Leu Lys Trp Gln Asp Glu Leu Asn Gln
 275 280 285

His Asn Arg Val Phe Phe Asp Ser Cys Asp Gly Phe Phe Thr Asn Tyr
 290 295 300

Asn Trp Arg Glu Glu His Leu Glu Arg Met Leu Gly Gln Ala Gly Glu
 305 310 315 320

Arg Arg Ala Asp Val Tyr Val Gly Val Asp Val Phe Ala Arg Gly Asn
 325 330 335

Val Val Gly Gly Arg Phe Asp Thr Asp Lys Ser Leu Glu Leu Ile Arg
 340 345 350

170

Lys His Gly Phe Ser Val Ala Leu Phe Ala Pro Gly Trp Val Tyr Glu
 355 360 365

Cys Leu Glu Lys Lys Asp Phe Phe Gln Asn Gln Asp Lys Phe Trp Gly
 370 375 380

Arg Leu Glu Arg Tyr Leu Pro Thr His Ser Ile Cys Ser Leu Pro Phe
 385 390 395 400

Val Thr Ser Phe Cys Leu Gly Met Gly Ala Arg Arg Val Cys Tyr Gly
 405 410 415

Gln Glu Glu Ala Val Gly Pro Trp Tyr His Leu Ser Ala Gln Glu Ile
 420 425 430

Gln Pro Leu Phe Gly Glu His Arg Leu Gly Gly Asp Gly Arg Gly Trp
 435 440 445

Val Arg Thr His Cys Cys Leu Glu Asp Ala Trp His Gly Gly Ser Ser
 450 455 460

Leu Leu Val Arg Gly Val Ile Pro Pro Glu Val Gly Asn Val Ala Val
 465 470 475 480

Arg Trp Val Ser Asp Gly Gly Arg Trp Ala His Gln Leu Leu Pro Ser
 485 490 495

His Val Val Ala Met Glu Leu Asp Arg Trp Gly Ser Gly Gly Gln Asn
 500 505 510

Lys Asp Arg Gly Gln Thr Gln Met Gly Phe Leu Lys Leu
 515 520 525

<210> 119
 <211> 327
 <212> PRT
 <213> Homo sapien

<400> 119

Met Phe Gln Pro Thr Lys Glu Ser Gly Leu Gly Gly Gly Leu Val Pro
 1 5 10 15

Trp Leu Arg Thr Gly Pro Arg Cys Gly Ser Ile Trp His Pro Gly Pro
 20 25 30

Leu Phe Leu Glu Gly Gly Ala Gly Gly Arg Asp Leu Glu Leu Ala Ser

171

35

40

45

Ile Ser Gln Trp Ser Leu His Gly Thr His His Arg Thr Phe Phe Pro
 50 55 60

Arg Leu Phe Ser Leu Gln Ala Pro Val Pro Pro Lys Ile Tyr Leu Ser
 65 70 75 80

Met Val Tyr Lys Leu Glu Gly Pro Thr Asp Val Thr Val Ala Leu Glu
 85 90 95

Leu Thr Thr Gly Asp Ala Gly Ser Cys His Ile Gly Gly Ile Ser Val
 100 105 110

Leu Asn Ala Glu Thr Ser Ser Arg His Ser Leu Arg Pro Leu Arg Val
 115 120 125

Pro Pro Thr Lys Leu Ala Arg Trp Val Gly Arg Cys Gly Arg Gln Leu
 130 135 140

Ser Gly Gly Trp Val Gln His Cys Tyr Glu Val Ser Leu Arg Gly Cys
 145 150 155 160

Leu Leu Leu Asp Leu Leu Val Cys Phe Ser Arg Pro Pro Gly Ser Arg
 165 170 175

Glu Glu Glu Ser Phe Thr Cys Arg Leu Gly Glu Ile Gln Val Val Asp
 180 185 190

Ala Ala Ser Leu Leu Ala Pro Leu Pro Gln Val Gln Ala Val Thr Ile
 195 200 205

Ser His Ile Arg Trp Gln Pro Ser Ala Ser Glu Arg Glu Gly Pro Pro
 210 215 220

Ala Leu Leu Gln Leu Ser Cys Thr Leu His Trp Ser Phe Leu Leu Ser
 225 230 235 240

Gln Val Arg Cys Phe Arg Ile His Cys Trp Gly Gly Met Ser Asp Asp
 245 250 255

Ser Pro Gly Arg Glu Leu Pro Arg Pro Glu Met Pro Met Phe Leu Gly
 260 265 270

Leu Ala Phe Ala Thr Gln Tyr Arg Ile Val Asp Leu Leu Val Glu Ala
 275 280 285

172

Ala Gly Pro Gly Gln Asp Arg Arg Met Glu Phe Leu Val Glu Pro Val
 290 295 300

Pro Lys Glu Gly Phe Arg Val Pro Gln Ala Glu Trp Gly Arg Ala Val
 305 310 315 320

Leu Leu Tyr Ser Ala Pro Ala
 325

<210> 120
 <211> 384
 <212> PRT
 <213> Homo sapien

<400> 120

Gln Ile Pro Arg Thr Val Ser Ser Cys Arg Thr Gly Leu Ser Pro Leu
 1 5 10 15

His Ile Ser Pro Pro Ser Ser Pro Ser Pro Pro Lys Pro Pro Leu Tyr
 20 25 30

Ser Ala Ser Val Ser Leu Asp Thr Leu Asp Ala Pro Tyr Glu Gly Ile
 35 40 45

Pro Tyr Gly Ile Ser Glu Leu Arg Cys Phe Ser Pro Gln Lys Asn Leu
 50 55 60

Ala Leu Gly Glu Asp Leu Ser Pro Gly Tyr Gly Gln Asp His Asp Val
 65 70 75 80

Gly Ala Phe Gly Thr Gln Ala Pro Cys Ser Trp Arg Glu Gly Leu Val
 85 90 95

Asp Ala Ile Trp Ser Trp Leu Arg Phe Leu Ser Gly Leu Ser Thr Ala
 100 105 110

Pro Ile Thr Gly Pro Phe Ser Pro Gly Tyr Phe Tyr Ser Leu Gln Ala
 115 120 125

Pro Val Pro Pro Lys Ile Tyr Leu Ser Met Val Tyr Lys Leu Glu Gly
 130 135 140

Pro Thr Asp Val Thr Val Ala Leu Glu Leu Thr Thr Gly Asp Ala Gly
 145 150 155 160

173

Ser Cys His Ile Gly Gly Ile Ser Val Leu Asn Ala Glu Thr Ser Ser
 165 170 175

Arg His Ser Leu Arg Pro Leu Arg Val Pro Pro Thr Lys Leu Ala Arg
 180 185 190

Trp Val Gly Arg Cys Gly Arg Gln Leu Ser Gly Gly Trp Val Gln His
 195 200 205

Cys Tyr Glu Val Ser Leu Arg Gly Cys Leu Leu Leu Asp Leu Leu Val
 210 215 220

Cys Phe Ser Arg Pro Pro Gly Ser Arg Glu Glu Glu Ser Phe Thr Cys
 225 230 235 240

Arg Leu Gly Glu Ile Gln Val Val Asp Ala Ala Ser Leu Leu Ala Pro
 245 250 255

Leu Pro Gln Val Gln Ala Val Thr Ile Ser His Ile Arg Trp Gln Pro
 260 265 270

Ser Ala Ser Glu Arg Glu Gly Pro Pro Ala Leu Leu Gln Leu Ser Cys
 275 280 285

Thr Leu His Trp Ser Phe Leu Leu Ser Gln Val Arg Cys Phe Arg Ile
 290 295 300

His Cys Trp Gly Gly Met Ser Asp Asp Ser Pro Gly Arg Glu Leu Pro
 305 310 315 320

Arg Pro Glu Met Pro Met Phe Leu Gly Leu Ala Phe Ala Thr Gln Tyr
 325 330 335

Arg Ile Val Asp Leu Leu Val Glu Ala Ala Gly Pro Gly Gln Asp Arg
 340 345 350

Arg Met Glu Phe Leu Val Glu Pro Val Pro Lys Glu Gly Phe Arg Val
 355 360 365

Pro Gln Ala Glu Trp Gly Arg Ala Val Leu Leu Tyr Ser Ala Pro Ala
 370 375 380

<210> 121
 <211> 450
 <212> PRT
 <213> Homo sapien

174

<400> 121

Gln Ile Pro Arg Thr Val Ser Ser Cys Arg Thr Gly Leu Ser Pro Leu
 1 5 10 15

His Ile Ser Pro Pro Ser Ser Pro Ser Pro Pro Lys Pro Pro Leu Tyr
 20 25 30

Ser Ala Ser Val Ser Leu Asp Thr Leu Asp Ala Pro Tyr Glu Gly Ile
 35 40 45

Pro Tyr Gly Ile Ser Glu Leu Arg Cys Phe Ser Pro Gln Lys Asn Leu
 50 55 60

Ala Leu Gly Glu Asp Leu Ser Pro Gly Tyr Gly Gln Asp His Asp Val
 65 70 75 80

Gly Ala Phe Gly Thr Gln Ala Pro Cys Ser Trp Arg Glu Gly Leu Val
 85 90 95

Asp Ala Ile Trp Ser Trp Leu Arg Phe Leu Ser Gly Leu Ser Thr Ala
 100 105 110

Pro Ile Thr Gly Pro Phe Ser Pro Gly Tyr Phe Tyr Ser Leu Gln Ala
 115 120 125

Pro Val Pro Pro Lys Ile Tyr Leu Ser Met Val Tyr Lys Leu Glu Gly
 130 135 140

Pro Thr Asp Val Thr Val Ala Leu Glu Leu Thr Thr Gly Asp Ala Gly
 145 150 155 160

Ser Cys His Ile Gly Gly Ile Ser Val Leu Asn Ala Glu Thr Ser Ser
 165 170 175

Arg His Ser Leu Arg Pro Leu Arg Val Pro Pro Thr Lys Leu Ala Arg
 180 185 190

Trp Val Gly Arg Cys Gly Arg Gln Leu Ser Gly Gly Trp Val Gln His
 195 200 205

Cys Tyr Glu Val Ser Leu Arg Gly Cys Leu Leu Leu Asp Leu Leu Val
 210 215 220

Cys Phe Ser Arg Pro Pro Gly Ser Arg Glu Glu Glu Ser Phe Thr Cys
 225 230 235 240

[illegible]

```
<210> 122
<211> 302
<212> PRT
<213> Homo sapien
```

176

<400> 122

Met Glu Ala Ala Ala Val Thr Val Thr Arg Ser Ala Thr Arg Arg Arg
 1 5 10 15

Arg Arg Gln Leu Gln Gly Leu Ala Ala Pro Glu Ala Gly Thr Gln Glu
 20 25 30

Glu Gln Glu Asp Gln Glu Pro Arg Pro Arg Arg Arg Arg Pro Gly Arg
 35 40 45

Ser Ile Lys Asp Glu Glu Glu Glu Thr Val Phe Arg Glu Val Val Ser
 50 55 60

Phe Ser Pro Asp Pro Leu Pro Val Arg Tyr Tyr Asp Lys Asp Thr Thr
 65 70 75 80

Lys Pro Ile Ser Phe Tyr Leu Ser Ser Leu Glu Glu Leu Leu Ala Trp
 85 90 95

Lys Pro Arg Leu Glu Asp Gly Phe Asn Val Ala Leu Glu Pro Leu Ala
 100 105 110

Cys Arg Gln Pro Pro Leu Ser Ser Gln Arg Pro Arg Thr Leu Leu Cys
 115 120 125

His Asp Met Met Gly Gly Tyr Leu Asp Asp Arg Phe Ile Gln Gly Ser
 130 135 140

Val Val Gln Thr Pro Tyr Ala Phe Tyr His Trp Gln Cys Ile Asp Val
 145 150 155 160

Phe Val Tyr Phe Ser His His Thr Val Thr Ile Pro Pro Val Gly Trp
 165 170 175

Thr Asn Thr Ala His Arg His Gly Val Cys Val Leu Gly Thr Phe Ile
 180 185 190

Thr Glu Trp Asn Glu Gly Gly Arg Leu Cys Glu Ala Phe Leu Ala Gly
 195 200 205

Asp Glu Arg Ser Tyr Gln Ala Val Ala Asp Arg Leu Val Gln Ile Thr
 210 215 220

Gln Phe Phe Arg Phe Asp Gly Trp Leu Ile Asn Ile Glu Asn Ser Leu
 225 230 235 240

Ser Leu Ala Ala Val Gly Asn Met Pro Pro Phe Leu Arg Tyr Leu Thr
245 250 255

Thr Gln Leu Leu Val Glu Ala Ala Gly Pro Gly Gln Asp Arg Arg Met
260 265 270

Glu Phe Leu Val Glu Pro Val Pro Lys Glu Gly Phe Arg Val Pro Gln
275 280 285

Ala Glu Trp Gly Arg Ala Val Leu Leu Tyr Ser Ala Pro Ala
290 295 300

```
<210> 123
<211> 162
<212> PRT
<213> Homo sapien
```

<400> 123

Met Pro Arg Trp Tyr Phe Leu Leu Ala Arg Cys Phe Gly Cys Ala Val
1 5 10 15

Ile Glu Asp Thr Trp His Tyr Phe Leu His Arg Leu Leu His His Lys
20 25 30

Arg Ile Tyr Lys Tyr Ile His Lys Val His His Glu Phe Gln Ala Pro
35 40 45

Phe Gly Met Glu Ala Glu Tyr Ala His Pro Leu Glu Thr Leu Ile Leu
50 55 60

Gly Thr Gly Phe Phe Ile Gly Ile Val Leu Leu Cys Asp His Val Ile
65 70 75 80

Leu Leu Trp Ala Trp Val Thr Ile Arg Leu Leu Glu Thr Ile Asp Val
85 90 95

His Ser Gly Tyr Asp Ile Pro Leu Asn Pro Leu Asn Leu Ile Pro Phe
100 105 110

Tyr Ala Gly Ser Arg His His Asp Phe His His Met Asn Phe Ile Gly
115 120 125

Asn Tyr Ala Ser Thr Phe Thr Trp Trp Asp Arg Ile Phe Gly Thr Asp
130 135 140

178

Ser Gln Tyr Asn Ala Tyr Asn Glu Lys Arg Lys Lys Phe Glu Lys Lys
 145 150 155 160

Thr Glu

<210> 124
 <211> 206
 <212> PRT
 <213> Homo sapien

<400> 124

Met Gly Glu Pro Gln Gly Ser Met Arg Ile Leu Val Thr Gly Gly Ser
 1 5 10 15

Gly Leu Val Gly Lys Ala Ile Gln Lys Val Val Ala Asp Gly Ala Gly
 20 25 30

Leu Pro Gly Glu Asp Trp Val Phe Val Ser Ser Lys Asp Ala Asp Leu
 35 40 45

Thr Asp Thr Ala Gln Thr Arg Ala Leu Phe Glu Lys Val Gln Pro Thr
 50 55 60

His Val Ile His Leu Ala Ala Met Val Gly Gly Leu Phe Arg Asn Ile
 65 70 75 80

Lys Tyr Asn Leu Asp Phe Trp Arg Lys Asn Val His Met Asn Asp Asn
 85 90 95

Val Leu His Ser Ala Phe Glu Val Gly Ala Arg Lys Val Val Ser Cys
 100 105 110

Leu Ser Thr Cys Ile Phe Pro Asp Lys Thr Thr Tyr Pro Ile Asp Glu
 115 120 125

Thr Met Ile His Asn Gly Pro Pro His Asn Ser Asn Phe Gly Tyr Ser
 130 135 140

Tyr Ala Lys Arg Met Ile Asp Val Gln Asn Ser Ala Gly Pro Thr Ser
 145 150 155 160

Ser Ser Thr Ala Ala Pro Ser Pro Leu Ser Ser Pro Pro Thr Ser Ser
 165 170 175

Gly Pro Thr Thr Thr Ser Thr Ser Arg Met Ala Thr Cys Cys Leu Ala
 180 185 190

179

Ser Ser Thr Arg Cys Thr Trp Pro Arg Ala Ala Ala Arg Pro
 195 200 205

<210> 125
 <211> 380
 <212> PRT
 <213> Homo sapien

<400> 125

Leu Gln Val Pro Ala Val Pro Gly Thr Leu Arg Ala Pro Gly Thr Pro
 1 5 10 15

Phe Pro Arg Val Pro Arg Pro Ser Leu Pro Ala Pro Pro Thr Trp
 20 25 30

Leu Arg Gly Gln Pro Glu Arg Thr Arg Pro Glu Ala Ala Val Gly Glu
 35 40 45

Pro Ala Val Gly Leu Asp Ala Gly Ala Thr Asp Met Gly Glu Pro Gln
 50 55 60

Gly Ser Met Arg Ile Leu Val Thr Gly Gly Ser Gly Leu Val Gly Lys
 65 70 75 80

Ala Ile Gln Lys Val Val Ala Asp Gly Ala Gly Leu Pro Gly Glu Asp
 85 90 95

Trp Val Phe Val Ser Ser Lys Asp Ala Asp Leu Thr Asp Thr Ala Gln
 100 105 110

Thr Arg Ala Leu Phe Glu Lys Val Gln Pro Thr His Val Ile His Leu
 115 120 125

Ala Ala Met Val Gly Gly Leu Phe Arg Asn Ile Lys Tyr Asn Leu Asp
 130 135 140

Phe Trp Arg Lys Asn Val His Met Asn Asp Asn Val Leu His Ser Ala
 145 150 155 160

Phe Glu Val Gly Ala Arg Lys Val Val Ser Cys Leu Ser Thr Cys Ile
 165 170 175

Phe Pro Asp Lys Thr Thr Tyr Pro Ile Asp Glu Thr Met Ile His Asn
 180 185 190

180

Gly Pro Pro His Asn Ser Asn Phe Gly Tyr Ser Tyr Ala Lys Arg Met
 195 200 205

Ile Asp Val Gln Asn Arg Ala Tyr Phe Gln Gln Tyr Gly Cys Thr Phe
 210 215 220

Thr Ala Val Ile Pro Thr Asn Val Phe Gly Pro His Asp Asn Phe Asn
 225 230 235 240

Ile Glu Asp Gly His Val Leu Pro Gly Leu Ile His Lys Val His Leu
 245 250 255

Ala Lys Ser Ser Gly Ser Ala Leu Thr Val Trp Gly Thr Gly Asn Pro
 260 265 270

Arg Arg Gln Phe Ile Tyr Ser Leu Asp Leu Ala Gln Leu Phe Ile Trp
 275 280 285

Val Leu Arg Glu Tyr Asn Glu Val Glu Pro Ile Ile Leu Ser Val Gly
 290 295 300

Glu Glu Asp Glu Val Ser Ile Lys Glu Ala Ala Glu Ala Val Val Glu
 305 310 315 320

Ala Met Asp Phe His Gly Glu Val Thr Phe Asp Thr Thr Lys Ser Asp
 325 330 335

Gly Gln Phe Lys Lys Thr Ala Ser Asn Ser Lys Leu Arg Thr Tyr Leu
 340 345 350

Pro Asp Phe Arg Phe Thr Pro Phe Lys Gln Ala Val Lys Glu Thr Cys
 355 360 365

Ala Trp Phe Thr Asp Asn Tyr Glu Gln Ala Arg Lys
 370 375 380

<210> 126
 <211> 380
 <212> PRT
 <213> Homo sapien

<400> 126

Leu Gln Val Pro Ala Val Pro Gly Thr Leu Arg Ala Pro Gly Thr Pro
 1 5 10 15

Phe Pro Arg Val Pro Arg Pro Ser Leu Pro Ala Pro Pro Pro Thr Trp
 20 25 30

181

Leu Arg Gly Gln Pro Glu Arg Thr Arg Pro Glu Ala Ala Val Gly Glu
 35 40 45

Pro Ala Val Gly Leu Asp Ala Gly Ala Thr Asp Met Gly Glu Pro Gln
 50 55 60

Gly Ser Met Arg Ile Leu Val Thr Gly Gly Ser Gly Leu Val Gly Lys
 65 70 75 80

Ala Ile Gln Lys Val Val Ala Asp Gly Ala Gly Leu Pro Gly Glu Asp
 85 90 95

Trp Val Phe Val Ser Ser Lys Asp Ala Asp Leu Thr Asp Thr Ala Gln
 100 105 110

Thr Arg Ala Leu Phe Glu Lys Val Gln Pro Thr His Val Ile His Leu
 115 120 125

Ala Ala Met Val Gly Gly Leu Phe Arg Asn Ile Lys Tyr Asn Leu Asp
 130 135 140

Phe Trp Arg Lys Asn Val His Met Asn Asp Asn Val Leu His Ser Ala
 145 150 155 160

Phe Glu Val Gly Ala Arg Lys Val Val Ser Cys Leu Ser Thr Cys Ile
 165 170 175

Phe Pro Asp Lys Thr Thr Tyr Pro Ile Asp Glu Thr Met Ile His Asn
 180 185 190

Gly Pro Pro His Asn Ser Asn Phe Gly Tyr Ser Tyr Ala Lys Arg Met
 195 200 205

Ile Asp Val Gln Asn Arg Ala Tyr Phe Gln Gln Tyr Gly Cys Thr Phe
 210 215 220

Thr Ala Val Ile Pro Thr Asn Val Phe Gly Pro His Asp Asn Phe Asn
 225 230 235 240

Ile Glu Asp Gly His Val Leu Pro Gly Leu Ile His Lys Val His Leu
 245 250 255

Ala Lys Ser Ser Gly Ser Ala Leu Thr Val Trp Gly Thr Gly Asn Pro
 260 265 270

182

Arg Arg Gln Phe Ile Tyr Ser Leu Asp Leu Ala Gln Leu Phe Ile Trp
 275 280 285

Val Leu Arg Glu Tyr Asn Glu Val Glu Pro Ile Ile Leu Ser Val Gly
 290 295 300

Glu Glu Asp Glu Val Ser Ile Lys Glu Ala Ala Glu Ala Val Val Glu
 305 310 315 320

Ala Met Asp Phe His Gly Glu Val Thr Phe Asp Thr Thr Lys Ser Asp
 325 330 335

Gly Gln Phe Lys Lys Thr Ala Ser Asn Ser Lys Leu Arg Thr Tyr Leu
 340 345 350

Pro Asp Phe Arg Phe Thr Pro Phe Lys Gln Ala Val Lys Glu Thr Cys
 355 360 365

Ala Trp Phe Thr Asp Asn Tyr Glu Gln Ala Arg Lys
 370 375 380

<210> 127
 <211> 334
 <212> PRT
 <213> Homo sapien

<400> 127

Met Arg Ala Leu Ala Ala Asn Arg Val Asn Asp Leu Cys Gln Glu Pro
 1 5 10 15

Pro Ser Gln Gly Cys Leu Pro Pro Pro Leu Val Ser Gln Arg Gly Val
 20 25 30

Glu Cys Thr Phe Ser Arg Pro Ser Gly Glu Ser Trp Val Gly Thr Ser
 35 40 45

Cys Ser Gly Leu Gly Gly Ser Ser Gly Pro Leu Arg Arg Cys Arg Leu
 50 55 60

Arg Ala Pro Arg Gly Thr Gly Leu Arg Arg Gly Ser Ala Ser Val Gln
 65 70 75 80

Leu Gly Leu Ser Gly Cys Gln Trp Thr Met Pro His Ser Glu Gly Leu
 85 90 95

Thr Leu Cys Gln Leu Pro Gln Lys Ser Gly Ala Pro Lys Asp Glu Ser

183

100	105	110
Val Met Thr Ser Ala Ser His Cys Leu Thr Leu Gln Gly Ala Thr Asp 115 120 125		
Met Gly Glu Pro Gln Gly Ser Met Arg Ile Leu Val Thr Gly Gly Ser 130 135 140		
Gly Leu Val Gly Lys Ala Ile Gln Lys Val Val Ala Asp Gly Ala Gly 145 150 155 160		
Leu Pro Gly Glu Asp Trp Val Phe Val Ser Ser Lys Asp Ala Asp Leu 165 170 175		
Thr Asp Thr Ala Gln Thr Arg Ala Leu Phe Glu Lys Val Gln Pro Thr 180 185 190		
His Val Ile His Leu Ala Ala Met Val Gly Gly Leu Phe Arg Asn Ile 195 200 205		
Lys Tyr Asn Leu Asp Phe Trp Arg Lys Asn Val His Met Asn Asp Asn 210 215 220		
Val Leu His Ser Ala Phe Glu Val Gly Ala Arg Lys Val Val Ser Cys 225 230 235 240		
Leu Ser Thr Cys Ile Phe Pro Asp Lys Thr Thr Tyr Pro Ile Asp Glu 245 250 255		
Thr Met Ile His Asn Gly Pro Pro His Asn Ser Asn Phe Gly Tyr Ser 260 265 270		
Tyr Ala Lys Arg Met Ile Asp Val Gln Asn Ser Ala Gly Pro Thr Ser 275 280 285		
Ser Ser Thr Ala Ala Pro Ser Pro Leu Ser Ser Pro Pro Thr Ser Ser 290 295 300		
Gly Pro Thr Thr Thr Ser Thr Ser Arg Met Ala Thr Cys Cys Leu Ala 305 310 315 320		
Ser Ser Thr Arg Cys Thr Trp Pro Arg Ala Ala Ala Arg Pro 325 330		
<210> 128		
<211> 327		

184

<212> PRT

<213> Homo sapien

<400> 128

His Tyr Ser Ala Thr Asp Met Gly Glu Pro Gln Gly Ser Met Arg Ile
 1 5 10 15

Leu Val Thr Gly Gly Ser Gly Leu Val Gly Lys Ala Ile Gln Lys Val
 20 25 30

Val Ala Asp Gly Ala Gly Leu Pro Gly Glu Asp Trp Val Phe Val Ser
 35 40 45

Ser Lys Asp Ala Asp Leu Thr Asp Thr Ala Gln Thr Arg Ala Leu Phe
 50 55 60

Glu Lys Val Gln Pro Thr His Val Ile His Leu Ala Ala Met Val Gly
 65 70 75 80

Gly Leu Phe Arg Asn Ile Lys Tyr Asn Leu Asp Phe Trp Arg Lys Asn
 85 90 95

Val His Met Asn Asp Asn Val Leu His Ser Ala Phe Glu Val Gly Ala
 100 105 110

Arg Lys Val Val Ser Cys Leu Ser Thr Cys Ile Phe Pro Asp Lys Thr
 115 120 125

Thr Tyr Pro Ile Asp Glu Thr Met Ile His Asn Gly Pro Pro His Asn
 130 135 140

Ser Asn Phe Gly Tyr Ser Tyr Ala Lys Arg Met Ile Asp Val Gln Asn
 145 150 155 160

Arg Ala Tyr Phe Gln Gln Tyr Gly Cys Thr Phe Thr Ala Val Ile Pro
 165 170 175

Thr Asn Val Phe Gly Pro His Asp Asn Phe Asn Ile Glu Asp Gly His
 180 185 190

Val Leu Pro Gly Leu Ile His Lys Val His Leu Ala Lys Ser Ser Gly
 195 200 205

Ser Ala Leu Thr Val Trp Gly Thr Gly Asn Pro Arg Arg Gln Phe Ile
 210 215 220

185

Tyr Ser Leu Asp Leu Ala Gln Leu Phe Ile Trp Val Leu Arg Glu Tyr
 225 230 235 240

Asn Glu Val Glu Pro Ile Ile Leu Ser Val Gly Glu Glu Asp Glu Val
 245 250 255

Ser Ile Lys Glu Ala Ala Glu Ala Val Val Glu Ala Met Asp Phe His
 260 265 270

Gly Glu Val Thr Phe Asp Thr Thr Lys Ser Asp Gly Gln Phe Lys Lys
 275 280 285

Thr Ala Ser Asn Ser Lys Leu Arg Thr Tyr Leu Pro Asp Phe Arg Phe
 290 295 300

Thr Pro Phe Lys Gln Ala Val Lys Glu Thr Cys Ala Trp Phe Thr Asp
 305 310 315 320

Asn Tyr Glu Gln Ala Arg Lys
 325

<210> 129
 <211> 161
 <212> PRT
 <213> Homo sapien

<400> 129

Met Gly Glu Pro Gln Gly Ser Met Arg Ile Leu Val Thr Gly Gly Ser
 1 5 10 15

Gly Leu Val Gly Lys Ala Ile Gln Lys Val Val Ala Asp Gly Ala Gly
 20 25 30

Leu Pro Gly Glu Asp Trp Val Phe Val Ser Ser Lys Asp Ala Asp Leu
 35 40 45

Thr Asp Thr Ala Gln Thr Arg Ala Leu Phe Glu Lys Val Gln Pro Thr
 50 55 60

His Val Ile His Leu Ala Ala Met Val Gly Gly Leu Phe Arg Asn Ile
 65 70 75 80

Lys Tyr Asn Leu Asp Phe Trp Arg Lys Asn Val His Met Asn Asp Asn
 85 90 95

Val Leu His Ser Ala Phe Glu Val Gly Ala Arg Lys Val Val Ser Cys
 100 105 110

186

Leu Ser Thr Cys Ile Phe Pro Asp Lys Thr Thr Tyr Pro Ile Asp Glu
 115 120 125

Thr Met Ile His Asn Gly Pro Pro His Asn Ser Asn Phe Gly Tyr Ser
 130 135 140

Tyr Ala Lys Arg Met Ile Asp Val Gln Asn Arg Ser Ser Arg Pro Ser
 145 150 155 160

Cys

<210> 130
 <211> 326
 <212> PRT
 <213> Homo sapien

<400> 130

Leu Gln Val Pro Ala Val Pro Gly Thr Leu Arg Ala Pro Gly Thr Pro
 1 5 10 15

Phe Pro Arg Val Pro Arg Pro Ser Leu Pro Ala Pro Pro Thr Trp
 20 25 30

Leu Arg Gly Gln Pro Glu Arg Thr Arg Pro Glu Ala Ala Val Gly Glu
 35 40 45

Pro Ala Val Gly Leu Asp Ala Gly Ala Thr Asp Met Gly Glu Pro Gln
 50 55 60

Gly Ser Met Arg Ile Leu Val Thr Gly Gly Ser Gly Leu Val Gly Lys
 65 70 75 80

Ala Ile Gln Lys Val Val Ala Asp Gly Ala Gly Leu Pro Gly Glu Asp
 85 90 95

Trp Val Phe Val Ser Ser Lys Asp Ala Asp Leu Thr Asp Thr Ala Gln
 100 105 110

Thr Arg Ala Leu Phe Glu Lys Val Gln Pro Thr His Val Ile His Leu
 115 120 125

Ala Ala Met Val Gly Gly Leu Phe Arg Asn Ile Lys Tyr Asn Leu Asp
 130 135 140

187

Phe Trp Arg Lys Asn Val His Met Asn Asp Asn Val Leu His Ser Ala
 145 150 155 160

Phe Glu Val Gly Ala Arg Lys Val Val Ser Cys Leu Ser Thr Cys Ile
 165 170 175

Phe Pro Asp Lys Thr Thr Tyr Pro Ile Asp Glu Thr Met Ile His Asn
 180 185 190

Gly Pro Pro His Asn Ser Asn Phe Gly Tyr Ser Tyr Ala Lys Arg Met
 195 200 205

Ile Asp Val Gln Asn Arg Ala Tyr Phe Gln Gln Tyr Gly Cys Thr Phe
 210 215 220

Thr Ala Val Ile Pro Thr Asn Val Phe Gly Pro His Asp Asn Phe Asn
 225 230 235 240

Ile Glu Asp Gly His Val Leu Pro Gly Leu Ile His Lys Val His Leu
 245 250 255

Ala Lys Ser Ser Gly Ser Ala Leu Thr Val Trp Gly Thr Gly Asn Pro
 260 265 270

Arg Arg Gln Phe Ile Tyr Ser Leu Asp Leu Ala Gln Leu Phe Ile Trp
 275 280 285

Val Leu Arg Glu Tyr Asn Glu Val Glu Pro Ile Ile Leu Ser Gly Gly
 290 295 300

Tyr Leu Ser Pro Gln Pro Pro Ser Ser Met Val Gly Gln Asp Pro Arg
 305 310 315 320

Leu Ser Trp Glu Ala Gly
 325

<210> 131
 <211> 216
 <212> PRT
 <213> Homo sapien

<400> 131

Met Gln Val Arg Thr Asp Pro Arg Ser Arg Gln Cys Trp Pro Leu Glu
 1 5 10 15

His Arg Thr Trp Leu Thr Asp Ser His Ser Ser Cys Leu Phe Pro Leu
 20 25 30

188

Pro Leu Glu Gln Pro Ser Leu Leu Gln Ser Asn Pro Cys Pro Ser Phe
 35 40 45

Leu Pro Leu Ser Arg Ala Ala Pro Pro Ala His Leu Arg Pro Gly Pro
 50 55 60

Ser Tyr Leu Leu Pro Leu Leu Ser Cys Pro Ile Pro Val Val Arg Arg
 65 70 75 80

Glu Ser Thr Gly Gln Arg Pro Ser Ser Thr Cys Asp Leu Gly Glu Cys
 85 90 95

Gln Ala Ser Pro Arg Gly Pro Gly Pro Arg Gly Pro Gly Arg Leu Cys
 100 105 110

Cys Gly Gly Ser Arg Val Arg Thr Gly Ala Ala Ser Pro Leu Ala Val
 115 120 125

Cys Leu Cys Pro Leu His Trp Pro Leu Glu Ala Gln Arg Pro Ser Gly
 130 135 140

Tyr Phe Pro Ser Ser Gly Leu Pro Leu Met Leu Phe Pro Ala Pro Phe
 145 150 155 160

Phe Tyr Leu Glu Thr Pro Ile Pro Ser His Pro Leu Gln Arg Ser Ser
 165 170 175

Gln Ser Cys Pro Gln His Gly Ser Leu His Ser Pro Trp Val Ser Pro
 180 185 190

Pro Val Ser Cys Leu Pro Arg Thr Pro Asp Leu Pro Leu Pro Gly Trp
 195 200 205

Pro Arg Trp Ile Leu Tyr Ser Asp
 210 215

<210> 132
 <211> 108
 <212> PRT
 <213> Homo sapien

<400> 132

Met Ala His Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu
 1 5 10 15

189

Arg Val Ala Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala
20 25 30

Ala Gly Gly Ser Arg Arg Pro Gly Val Pro Gly Pro Asp Ala Ala Gly
35 40 45

Val Gly Ala Pro Arg Arg Thr Ala Pro Leu Asn Gln Arg Val Tyr Ser
50 55 60

Ser Leu Gly Ala Ser Val Val Thr Glu Leu Arg Cys Gln Cys Leu Gln
65 70 75 80

Thr Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Asn Val Arg
85 90 95

Ser Pro Gly Pro His Cys Ala Gln Thr Glu Val Met
100 105

<210> 133
<211> 142
<212> PRT
<213> Homo sapien

<400> 133

Lys Gly Ser Pro Ile Leu Gly Ser His Thr Ala Arg Val Ala Gly Thr
1 5 10 15

Ser Pro Pro Ala Leu Pro Leu Leu Ala Gln Leu Pro Asp Ala Ser Ala
20 25 30

Glu Pro His Gly Pro Arg His Ala Leu Arg Arg Pro Gln Gln Ser Pro
35 40 45

Ala Pro Ala Gly Gly Ala Ala Ala Pro Ala Pro Gly Gly Arg Gln Pro
50 55 60

Ala Arg Ser Arg Trp Val Pro Ala Pro Trp Gly Pro Arg Ala Gly Arg
65 70 75 80

Gly Trp Gly Gly Arg Pro Ala Pro Thr Ala Pro Leu Asn Gln Arg Val
85 90 95

Tyr Ser Ser Leu Gly Ala Ser Val Val Thr Glu Leu Arg Cys Gln Cys
100 105 110

Leu Gln Thr Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Asn
115 120 125

190

Val Arg Ser Pro Gly Pro His Cys Ala Gln Thr Glu Val Met
 130 135 140

<210> 134
 <211> 482
 <212> PRT
 <213> Homo sapien

<400> 134

Met Val Met Glu Lys Pro Ser Pro Leu Leu Val Gly Arg Glu Phe Val
 1 5 10 15

Arg Gln Tyr Tyr Thr Leu Leu Asn Lys Ala Pro Glu Tyr Leu His Arg
 20 25 30

Phe Tyr Gly Arg Asn Ser Ser Tyr Val His Gly Gly Val Asp Ala Ser
 35 40 45

Gly Lys Pro Gln Glu Ala Val Tyr Gly Gln Asn Asp Ile His His Lys
 50 55 60

Val Leu Ser Leu Asn Phe Ser Glu Cys His Thr Lys Ile Arg His Val
 65 70 75 80

Asp Ala His Ala Thr Leu Ser Asp Gly Val Val Val Gln Val Met Gly
 85 90 95

Leu Leu Ser Asn Ser Gly Gln Pro Glu Arg Lys Phe Met Gln Thr Phe
 100 105 110

Val Leu Ala Pro Glu Gly Ser Val Pro Asn Lys Phe Tyr Val His Asn
 115 120 125

Asp Met Phe Arg Tyr Glu Asp Glu Val Phe Gly Asp Ser Glu Pro Glu
 130 135 140

Leu Asp Glu Glu Ser Glu Asp Glu Val Glu Glu Glu Gln Glu Glu Arg
 145 150 155 160

Gln Pro Ser Pro Glu Pro Val Gln Glu Asn Ala Asn Ser Gly Tyr Tyr
 165 170 175

Glu Ala His Pro Val Thr Asn Gly Ile Glu Glu Pro Leu Glu Glu Ser
 180 185 190

191

Ser His Glu Pro Glu Pro Glu Pro Glu Ser Glu Thr Lys Thr Glu Glu
 195 200 205

Leu Lys Pro Gln Val Glu Glu Lys Asn Leu Glu Glu Leu Glu Glu Lys
 210 215 220

Ser Thr Thr Pro Pro Pro Ala Glu Pro Val Ser Leu Pro Gln Glu Pro
 225 230 235 240

Pro Lys Ala Phe Ser Trp Ala Ser Val Thr Ser Lys Asn Leu Pro Pro
 245 250 255

Ser Gly Thr Val Ser Ser Ser Gly Ile His Pro Met Leu Lys His Gln
 260 265 270

Ser His Ser Gln Glu Ser Lys Leu Asn Gln Lys Phe Asn Leu Ser His
 275 280 285

Leu Val Cys Val Asn Asn Asp Leu Glu Asn Asp Leu Val Phe Leu Leu
 290 295 300

Glu Asp Gln Asp Gln Ala Glu Glu Ile Met Glu Gln Asn Asp Ser Asp
 305 310 315 320

Asn Arg Arg Ile Ile Arg Tyr Pro Asp Ser His Gln Leu Phe Val Gly
 325 330 335

Asn Leu Pro His Asp Ile Asp Glu Asn Glu Leu Lys Glu Phe Phe Met
 340 345 350

Ser Phe Gly Asn Val Val Glu Leu Arg Ile Asn Thr Lys Gly Val Gly
 355 360 365

Gly Lys Leu Pro Asn Phe Gly Phe Val Val Phe Asp Asp Ser Glu Pro
 370 375 380

Val Gln Arg Ile Leu Ile Ala Lys Pro Ile Met Phe Arg Gly Glu Val
 385 390 395 400

Arg Leu Asn Val Glu Glu Lys Lys Thr Arg Ala Ala Arg Glu Arg Glu
 405 410 415

Thr Arg Gly Gly Gly Asp Asp Arg Arg Asp Ile Arg Arg Asn Asp Arg
 420 425 430

Gly Pro Gly Gly Pro Arg Gly Ile Val Gly Gly Gly Met Met Arg Asp

192

435

440

445

Arg Asp Gly Arg Gly Pro Pro Pro Arg Gly Gly Met Ala Gln Lys Leu
 450 455 460

Gly Ser Gly Arg Gly Thr Gly Gln Met Glu Gly Arg Phe Thr Gly Gln
 465 470 475 480

Arg Arg

<210> 135
 <211> 392
 <212> PRT
 <213> Homo sapien

<400> 135

Leu Ser Arg Ser Trp Val Cys Cys Leu Thr Val Asp Asn Gln Lys Glu
 1 5 10 15

Ser Leu Cys Lys Pro Leu Phe Trp Leu Leu Lys Asp Leu Phe Gln Ile
 20 25 30

Asn Phe Met Phe Thr Met Ile Cys Phe Val Met Lys Met Lys Cys Tyr
 35 40 45

Gly Asp Ser Glu Pro Glu Leu Asp Glu Glu Ser Glu Asp Glu Val Glu
 50 55 60

Glu Glu Gln Glu Glu Arg Gln Pro Ser Pro Glu Pro Val Gln Glu Asn
 65 70 75 80

Ala Asn Ser Gly Tyr Tyr Glu Ala His Pro Val Thr Asn Gly Ile Glu
 85 90 95

Glu Pro Leu Glu Glu Ser Ser His Glu Pro Glu Pro Glu Pro Glu Ser
 100 105 110

Glu Thr Lys Thr Glu Glu Leu Lys Pro Gln Val Glu Glu Lys Asn Leu
 115 120 125

Glu Glu Leu Glu Glu Lys Ser Thr Thr Pro Pro Pro Ala Glu Pro Val
 130 135 140

Ser Leu Pro Gln Glu Pro Pro Lys Ala Phe Ser Trp Ala Ser Val Thr
 145 150 155 160

193

Ser Lys Asn Leu Pro Pro Ser Gly Thr Val Ser Ser Ser Gly Ile Pro
 165 170 175

Pro His Val Lys Ala Pro Val Ser Gln Pro Arg Val Glu Ala Lys Pro
 180 185 190

Glu Val Gln Ser Gln Pro Pro Arg Val Arg Glu Gln Arg Pro Arg Glu
 195 200 205

Arg Pro Gly Phe Pro Pro Arg Gly Pro Arg Pro Gly Arg Gly Asp Met
 210 215 220

Glu Gln Asn Asp Ser Asp Asn Arg Arg Ile Ile Arg Tyr Pro Asp Ser
 225 230 235 240

His Gln Leu Phe Val Gly Asn Leu Pro His Asp Ile Asp Glu Asn Glu
 245 250 255

Leu Lys Glu Phe Phe Met Ser Phe Gly Asn Val Val Glu Leu Arg Ile
 260 265 270

Asn Thr Lys Gly Val Gly Gly Lys Leu Pro Asn Phe Gly Phe Val Val
 275 280 285

Phe Asp Asp Ser Glu Pro Val Gln Arg Ile Leu Ile Ala Lys Pro Ile
 290 295 300

Met Phe Arg Gly Glu Val Arg Leu Asn Val Glu Glu Lys Lys Thr Arg
 305 310 315 320

Ala Ala Arg Glu Arg Glu Thr Arg Gly Gly Gly Asp Asp Arg Arg Asp
 325 330 335

Ile Arg Arg Asn Asp Arg Gly Pro Gly Gly Pro Arg Gly Ile Val Gly
 340 345 350

Gly Gly Met Met Arg Asp Arg Asp Gly Arg Gly Pro Pro Pro Arg Gly
 355 360 365

Gly Met Ala Gln Lys Leu Gly Ser Gly Arg Gly Thr Gly Gln Met Glu
 370 375 380

Gly Arg Phe Thr Gly Gln Arg Arg
 385 390

194

<210> 136
 <211> 316
 <212> PRT
 <213> Homo sapien

<400> 136

Asp Trp Glu Glu Lys Arg Val Leu Ala Ile Cys Leu Ala Ser Gln Ser
 1 5 10 15

Glu Asp Glu Val Glu Glu Glu Gln Glu Glu Arg Gln Pro Ser Pro Glu
 20 25 30

Pro Val Gln Glu Asn Ala Asn Ser Gly Tyr Tyr Glu Ala His Pro Val
 35 40 45

Thr Asn Gly Ile Glu Glu Pro Leu Glu Glu Ser Ser His Glu Pro Glu
 50 55 60

Pro Glu Pro Glu Ser Glu Thr Lys Thr Glu Glu Leu Lys Pro Gln Val
 65 70 75 80

Glu Glu Lys Asn Leu Glu Glu Leu Glu Glu Lys Ser Thr Thr Pro Pro
 85 90 95

Pro Ala Glu Pro Val Ser Leu Pro Gln Glu Pro Pro Lys Pro Arg Val
 100 105 110

Glu Ala Lys Pro Glu Val Gln Ser Gln Pro Pro Arg Val Arg Glu Gln
 115 120 125

Arg Pro Arg Glu Arg Pro Gly Phe Pro Pro Arg Gly Pro Arg Pro Gly
 130 135 140

Arg Gly Asp Met Glu Gln Asn Asp Ser Asp Asn Arg Arg Ile Ile Arg
 145 150 155 160

Tyr Pro Asp Ser His Gln Leu Phe Val Gly Asn Leu Pro His Asp Ile
 165 170 175

Asp Glu Asn Glu Leu Lys Glu Phe Phe Met Ser Phe Gly Asn Val Val
 180 185 190

Glu Leu Arg Ile Asn Thr Lys Gly Val Gly Gly Lys Leu Pro Asn Phe
 195 200 205

Gly Phe Val Val Phe Asp Asp Ser Glu Pro Val Gln Arg Ile Leu Ile
 210 215 220

195

Ala Lys Pro Ile Met Phe Arg Gly Glu Val Arg Leu Asn Val Glu Glu
 225 230 235 240

Lys Lys Thr Arg Ala Ala Arg Glu Arg Glu Thr Arg Gly Gly Gly Asp
 245 250 255

Asp Arg Arg Asp Ile Arg Arg Asn Asp Arg Gly Pro Gly Gly Pro Arg
 260 265 270

Gly Ile Val Gly Gly Gly Met Met Arg Asp Arg Asp Gly Arg Gly Pro
 275 280 285

Pro Pro Arg Gly Gly Met Ala Gln Lys Leu Gly Ser Gly Arg Gly Thr
 290 295 300

Gly Gln Met Glu Gly Arg Phe Thr Gly Gln Arg Arg
 305 310 315

<210> 137

<211> 314

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> X=any amino acid

<400> 137

Leu Gly Gly Xaa Glu Ser Gln Leu Leu Leu Ala Ser Gln Ser Glu Asp
 1 5 10 15

Glu Val Glu Glu Glu Gln Glu Glu Arg Gln Pro Ser Pro Glu Pro Val
 20 25 30

Gln Glu Asn Ala Asn Ser Gly Tyr Tyr Glu Ala His Pro Val Thr Asn
 35 40 45

Gly Ile Glu Glu Pro Leu Glu Glu Ser Ser His Glu Pro Glu Pro Glu
 50 55 60

Pro Glu Ser Glu Thr Lys Thr Glu Glu Leu Lys Pro Gln Val Glu Glu
 65 70 75 80

Lys Asn Leu Glu Glu Leu Glu Glu Lys Ser Thr Thr Pro Pro Pro Ala
 85 90 95

196

Glu Pro Val Ser Leu Pro Gln Glu Pro Pro Lys Pro Arg Val Glu Ala
 100 105 110

Lys Pro Glu Val Gln Ser Gln Pro Pro Arg Val Arg Glu Gln Arg Pro
 115 120 125

Arg Glu Arg Pro Gly Phe Pro Pro Arg Gly Pro Arg Pro Gly Arg Gly
 130 135 140

Asp Met Glu Gln Asn Asp Ser Asp Asn Arg Arg Ile Ile Arg Tyr Pro
 145 150 155 160

Asp Ser His Gln Leu Phe Val Gly Asn Leu Pro His Asp Ile Asp Glu
 165 170 175

Asn Glu Leu Lys Glu Phe Phe Met Ser Phe Gly Asn Val Val Glu Leu
 180 185 190

Arg Ile Asn Thr Lys Gly Val Gly Gly Lys Leu Pro Asn Phe Gly Phe
 195 200 205

Val Val Phe Asp Asp Ser Glu Pro Val Gln Arg Ile Leu Ile Ala Lys
 210 215 220

Pro Ile Met Phe Arg Gly Glu Val Arg Leu Asn Val Glu Glu Lys Lys
 225 230 235 240

Thr Arg Ala Ala Arg Glu Arg Glu Thr Arg Gly Gly Gly Asp Asp Arg
 245 250 255

Arg Asp Ile Arg Arg Asn Asp Arg Gly Pro Gly Gly Pro Arg Gly Ile
 260 265 270

Val Gly Gly Gly Met Met Arg Asp Arg Asp Gly Arg Gly Pro Pro Pro
 275 280 285

Arg Gly Gly Met Ala Gln Lys Leu Gly Ser Gly Arg Gly Thr Gly Gln
 290 295 300

Met Glu Gly Arg Phe Thr Gly Gln Arg Arg
 305 310

<210> 138
 <211> 169
 <212> PRT

197

<213> Homo sapien

<400> 138

Met Glu Gln Asn Asp Ser Asp Asn Arg Arg Ile Ile Arg Tyr Pro Asp
 1 5 10 15

Ser His Gln Leu Phe Val Gly Asn Leu Pro His Asp Ile Asp Glu Asn
 20 25 30

Glu Leu Lys Glu Phe Phe Met Ser Phe Gly Asn Val Val Glu Leu Arg
 35 40 45

Ile Asn Thr Lys Gly Val Gly Gly Lys Leu Pro Asn Phe Gly Phe Val
 50 55 60

Val Phe Asp Asp Ser Glu Pro Val Gln Arg Ile Leu Ile Ala Lys Pro
 65 70 75 80

Ile Met Phe Arg Gly Glu Val Arg Leu Asn Val Glu Glu Lys Lys Thr
 85 90 95

Arg Ala Ala Arg Glu Arg Glu Thr Arg Gly Gly Gly Asp Asp Arg Arg
 100 105 110

Asp Ile Arg Arg Asn Asp Arg Gly Pro Gly Gly Pro Arg Gly Ile Val
 115 120 125

Gly Gly Gly Met Met Arg Asp Arg Asp Gly Arg Gly Pro Pro Pro Arg
 130 135 140

Gly Gly Met Ala Gln Lys Leu Gly Ser Gly Arg Gly Thr Gly Gln Met
 145 150 155 160

Glu Gly Arg Phe Thr Gly Gln Arg Arg
 165

<210> 139

<211> 147

<212> PRT

<213> Homo sapien

<400> 139

Met Gly Arg Val Arg Thr Lys Thr Val Lys Lys Ala Ala Arg Val Ile
 1 5 10 15

Ile Glu Lys Tyr Tyr Thr Arg Leu Gly Asn Asp Phe His Thr Asn Lys
 20 25 30

Arg Val Cys Glu Glu Ile Ala Ile Ile Pro Ser Lys Lys Leu Arg Asn
35 40 45

Lys Ile Ala Gly Tyr Val Thr His Leu Met Lys Arg Ile Gln Arg Gly
50 55 60

Pro Val Arg Gly Ile Ser Ile Lys Leu Gln Glu Glu Glu Arg Glu Arg
65 70 75 80

Arg Asp Asn Tyr Val Pro Glu Val Ser Ala Leu Asp Gln Glu Ile Ile
85 90 95

Glu Val Asp Pro Asp Thr Lys Glu Met Leu Lys Leu Leu Asp Phe Gly
100 105 110

Ser Leu Ser Asn Leu Gln Val Ile His Pro Asn Cys Arg Leu Ser Asp
115 120 125

Leu Lys Val Gly Gln Thr Ala Val Gly Met Asn Phe Lys Thr Pro Arg
130 135 140

Gly Pro Val
145

```
<210> 140
<211> 166
<212> PRT
<213> Homo sapien
```

```
<220> .
<221> MISC_FEATURE
<222> (129)..(129)
<223> X=any amino acid
```

```

<220>
<221> MISC_FEATURE
<222> (134)..(134)
<223> X=any amino acid

```

<400> 140

Ala Leu Thr Gly Phe Ala Cys Ala Ser Cys Phe Leu Phe Tyr Gln Gly
1 5 10 15

Pro Ala Asn Met Gly Arg Val Arg Thr Lys Thr Val Lys Lys Ala Ala
20 25 30

199

Arg Val Ile Ile Glu Lys Tyr Tyr Thr Arg Leu Gly Asn Asp Phe His
 35 40 45

Thr Asn Lys Arg Val Cys Glu Glu Ile Ala Ile Ile Pro Ser Lys Lys
 50 55 60

Leu Arg Asn Lys Ile Ala Gly Tyr Val Thr His Leu Met Lys Arg Ile
 65 70 75 80

Gln Arg Gly Pro Val Arg Gly Ile Ser Ile Lys Leu Gln Glu Glu Glu
 85 90 95

Arg Glu Arg Arg Asp Asn Tyr Val Pro Glu Val Ser Ala Leu Asp Gln
 100 105 110

Glu Ile Ile Glu Val Asp Pro Asp Thr Lys Glu Met Leu Lys Leu Leu
 115 120 125

Xaa Phe Gly Ser Leu Xaa Asn Leu Gln Val Ile His Pro Asn Cys Arg
 130 135 140

Leu Ser Asp Leu Lys Val Gly Gln Thr Ala Val Gly Met Asn Phe Lys
 145 150 155 160

Thr Pro Arg Gly Pro Val
 165

<210> 141
 <211> 254
 <212> PRT
 <213> Homo sapien

<400> 141

Met Ser Val Asn Ala Ile Arg Lys Gln Ser Thr Asp Glu Glu Val Thr
 1 5 10 15

Ser Leu Ala Lys Ser Leu Ile Lys Ser Trp Lys Lys Leu Leu Asp Gly
 20 25 30

Pro Ser Thr Glu Lys Asp Leu Asp Glu Lys Lys Lys Glu Pro Ala Ile
 35 40 45

Thr Ser Gln Asn Ser Pro Glu Ala Arg Glu Glu Ser Thr Ser Ser Gly
 50 55 60

Asn Val Ser Asn Arg Lys Asp Glu Thr Asn Ala Arg Asp Thr Tyr Val
 65 70 75 80

200

Ser Ser Phe Pro Arg Ala Pro Ser Thr Ser Asp Ser Val Arg Leu Lys
85 90 95

Cys Arg Glu Met Leu Ala Ala Ala Leu Arg Thr Gly Asp Asp Tyr Ile
100 105 110

Ala Ile Gly Ala Asp Glu Glu Glu Leu Gly Ser Gln Ile Glu Glu Ala
115 120 125

Ile Tyr Gln Glu Ile Arg Asn Thr Asp Met Lys Tyr Lys Asn Arg Val
130 135 140

Arg Ser Arg Ile Ser Asn Leu Lys Asp Ala Lys Asn Pro Asn Leu Arg
145 150 155 160

Lys Asn Val Leu Cys Gly Asn Ile Pro Pro Asp Leu Phe Ala Arg Met
165 170 175

Thr Ala Glu Glu Met Ala Ser Asp Glu Leu Lys Glu Met Arg Lys Asn
180 185 190

Leu Thr Lys Glu Ala Ile Arg Glu His Gln Met Ala Lys Thr Gly Gly
195 200 205

Thr Gln Thr Asp Leu Phe Thr Cys Gly Lys Cys Lys Lys Lys Asn Cys
210 215 220

Thr Tyr Thr Gln Val Gln Thr Arg Ser Ala Asp Glu Pro Met Thr Thr
225 230 235 240

Phe Val Val Cys Asn Glu Cys Gly Asn Arg Trp Lys Phe Cys
245 250

<210> 142

<211> 302

<212> PRT

<213> Homo sapien

<400> 142

Arg Gly Leu Asn Val Arg Leu Val Ile Ser Thr Val Leu His Val Cys
1 5 10 15

Leu Ala Ile Lys Asn Ala Ala Gly Ala Leu Asp Leu Leu Lys Glu Leu
20 25 30

201

Lys Asn Ile Pro Met Thr Leu Glu Leu Leu Gln Ser Thr Arg Ile Gly
 35 40 45

Met Ser Val Asn Ala Ile Arg Lys Gln Ser Thr Asp Glu Glu Val Thr
 50 55 60

Ser Leu Ala Lys Ser Leu Ile Lys Ser Trp Lys Lys Leu Leu Asp Gly
 65 70 75 80

Pro Ser Thr Glu Lys Asp Leu Asp Glu Lys Lys Lys Glu Pro Ala Ile
 85 90 95

Thr Ser Gln Asn Ser Pro Glu Ala Arg Glu Glu Ser Thr Ser Ser Gly
 100 105 110

Asn Val Ser Asn Arg Lys Asp Glu Thr Asn Ala Arg Asp Thr Tyr Val
 115 120 125

Ser Ser Phe Pro Arg Ala Pro Ser Thr Ser Asp Ser Val Arg Leu Lys
 130 135 140

Cys Arg Glu Met Leu Ala Ala Ala Leu Arg Thr Gly Asp Asp Tyr Ile
 145 150 155 160

Ala Ile Gly Ala Asp Glu Glu Glu Leu Gly Ser Gln Ile Glu Glu Ala
 165 170 175

Ile Tyr Gln Glu Ile Arg Asn Thr Asp Met Lys Tyr Lys Asn Arg Val
 180 185 190

Arg Ser Arg Ile Ser Asn Leu Lys Asp Ala Lys Asn Pro Asn Leu Arg
 195 200 205

Lys Asn Val Leu Cys Gly Asn Ile Pro Pro Asp Leu Phe Ala Arg Met
 210 215 220

Thr Ala Glu Glu Met Ala Ser Asp Glu Leu Lys Glu Met Arg Lys Asn
 225 230 235 240

Leu Thr Lys Glu Ala Ile Arg Glu His Gln Met Ala Lys Thr Gly Gly
 245 250 255

Thr Gln Thr Asp Leu Phe Thr Cys Gly Lys Cys Lys Lys Lys Asn Cys
 260 265 270

Thr Tyr Thr Gln Val Gln Thr Arg Ser Ala Asp Glu Pro Met Thr Thr

202

275

280

285

Phe Val Val Cys Asn Glu Cys Gly Asn Arg Trp Lys Phe Cys
 290 295 300

<210> 143
 <211> 225
 <212> PRT
 <213> Homo sapien

<400> 143

Met Val Ser His Ser Glu Leu Arg Lys Leu Phe Tyr Ser Ala Asp Ala
 1 5 10 15

Val Cys Phe Asp Val Asp Ser Thr Val Ile Arg Glu Glu Gly Ile Asp
 20 25 30

Glu Leu Ala Lys Ile Cys Gly Val Glu Asp Ala Val Ser Glu Met Thr
 35 40 45

Arg Arg Ala Met Gly Gly Ala Val Pro Phe Lys Ala Ala Leu Thr Glu
 50 55 60

Arg Leu Ala Leu Ile Gln Pro Ser Arg Glu Gln Val Gln Arg Leu Ile
 65 70 75 80

Ala Glu Gln Pro Pro His Leu Thr Pro Gly Ile Arg Glu Leu Val Ser
 85 90 95

Arg Leu Gln Glu Arg Asn Val Gln Val Phe Leu Ile Ser Gly Gly Phe
 100 105 110

Arg Ser Ile Val Glu His Val Ala Ser Lys Leu Asn Ile Pro Ala Thr
 115 120 125

Asn Val Phe Ala Asn Arg Leu Lys Phe Tyr Phe Asn Gly Glu Tyr Ala
 130 135 140

Gly Phe Asp Glu Thr Gln Pro Thr Ala Glu Ser Gly Gly Lys Gly Lys
 145 150 155 160

Val Ile Lys Leu Leu Lys Glu Lys Phe His Phe Lys Lys Ile Ile Met
 165 170 175

Ile Gly Asp Gly Ala Thr Asp Met Glu Ala Cys Pro Pro Ala Asp Ala
 180 185 190

203

Phe Ile Gly Phe Gly Gly Asn Val Ile Arg Gln Gln Val Lys Asp Asn
 195 200 205

Ala Lys Trp Tyr Ile Thr Asp Phe Val Glu Leu Leu Gly Glu Leu Glu
 210 215 220

Glu
 225

<210> 144
 <211> 249
 <212> PRT
 <213> Homo sapien

<400> 144

Met Lys Gln Thr Asn Ile Gln Lys Asn Thr Asn Thr Arg Asp Thr Ser
 1 5 10 15

Lys Lys Thr Lys Asp Gln Leu Ile Ile Asp Ala Gly Gln Lys His Phe
 20 25 30

Gly Ala Thr Val Cys Lys Ser Cys Gly Met Ile Tyr Thr Ala Ser Asn
 35 40 45

Pro Glu Asp Glu Met Gln His Val Gln His His His Arg Phe Leu Glu
 50 55 60

Gly Ile Lys Tyr Val Gly Trp Lys Lys Glu Arg Val Val Ala Glu Phe
 65 70 75 80

Trp Asp Gly Lys Ile Val Leu Val Leu Pro His Asp Pro Ser Phe Ala
 85 90 95

Ile Lys Lys Val Glu Asp Val Gln Glu Leu Val Asp Asn Glu Leu Gly
 100 105 110

Phe Gln Gln Val Val Pro Lys Cys Pro Asn Lys Ile Lys Thr Phe Leu
 115 120 125

Phe Ile Ser Asp Glu Lys Arg Val Val Gly Cys Leu Ile Ala Glu Pro
 130 135 140

Ile Lys Gln Ala Phe Arg Val Leu Ser Glu Pro Ile Gly Pro Glu Ser
 145 150 155 160

Pro Ser Ser Thr Glu Cys Pro Arg Ala Trp Gln Cys Ser Asp Val Pro

204

165 170 175

Glu Pro Ala Val Cys Gly Ile Ser Arg Ile Trp Val Phe Arg Leu Lys
180 185 190

Arg Arg Lys Arg Ile Ala Arg Arg Leu Val Asp Thr Leu Arg Asn Cys
195 200 205

Phe Met Phe Gly Cys Phe Leu Ser Thr Asp Glu Ile Ala Phe Ser Asp
210 215 220

Pro Thr Pro Asp Gly Lys Leu Phe Ala Thr Lys Tyr Cys Asn Thr Pro
225 230 235 240

Asn Phe Leu Val Tyr Asn Phe Asn Ser
245

<210> 145
<211> 113
<212> PRT
<213> Homo sapien

<400> 145

Met Lys Ser Phe Ser Lys Ser Ser Asn Lys Cys Thr Leu Asn Thr Ser
1 5 10 15

Thr Val Arg Glu Phe Leu Ser Phe Arg Met Asn Ala Ile His Thr Lys
20 25 30

Glu Leu Leu Leu Thr Ser His Leu Gln Ser Pro Pro Gly His Arg Gln
35 40 45

Asp Pro Phe Asn Lys Ser Ser Ser Glu Thr Pro Ile Val Gln Asn Leu
50 55 60

Gln Leu Ala Thr Gly Tyr His His Ser Leu Trp Leu Cys Lys Ile Lys
65 70 75 80

Asp Leu Glu Glu Gly Trp Gly Gly Gly Ser Tyr Glu Lys Arg Gln Glu
85 90 95

Lys Ser Ser Phe Asp Pro Met Leu Ser Glu Ser Val His Glu Glu Glu
100 105 110

Ser

205

<210> 146
 <211> 102
 <212> PRT
 <213> Homo sapien

<400> 146

Met Val Thr Glu Glu Lys Arg Ser Glu Ala Arg Glu Asn Glu Arg Ser
 1 5 10 15

Leu Ala Phe Val Lys Met Val Gly His His Val Ala Phe Leu Glu Ala
 20 25 30

Asp Val Leu Gln Ala Glu Arg Asp His Gly Ala Phe Pro Gln Ala Leu
 35 40 45

Arg Arg Trp Leu Gly Ser Ala Gly Leu Pro Ser Phe Arg Asn Lys Ser
 50 55 60

Pro Ala Pro Val Pro Val Thr Tyr Glu Leu Pro Thr Leu Tyr Arg Thr
 65 70 75 80

Glu Asp Tyr Phe Pro Val Asp Ala Gly Glu Ala Gln His His Pro Arg
 85 90 95

Thr Cys Pro Arg Pro Leu
 100

<210> 147
 <211> 412
 <212> PRT
 <213> Homo sapien

<400> 147

Met Thr His Arg Arg Phe Lys Val Thr Ser Thr Val Ala Ala Ala Ser
 1 5 10 15

Leu Leu Pro Leu Gln Asp Glu Lys Glu Val Leu Leu Cys Lys Pro Ala
 20 25 30

Trp Leu Ser Pro Ser Gly Thr Arg Thr Gly Gly Phe Leu Ala Val Pro
 35 40 45

Gly Pro Pro Leu Arg Ala Lys Gly Pro Pro Val Leu Trp Pro Pro Pro
 50 55 60

Ala His Pro Pro Arg Val Pro Gly Arg Glu His Ser Arg Trp Gly Arg
 65 70 75 80

206

Ser Pro Pro Ala Gln Arg Ala Ala Leu Gly Leu Arg Pro Tyr Leu Leu
 85 90 95

Leu Leu Leu Pro Pro Ala Gln Leu Phe Asn Val Tyr Pro Trp Leu Gly
 100 105 110

Ala Leu Leu Gln Leu His Arg Pro Val Leu Arg Lys Ile Glu Glu Val
 115 120 125

Arg Ala Ile Leu Arg Thr Leu Leu Glu Ala Arg Arg Pro His Val Cys
 130 135 140

Pro Gly Asp Pro Val Cys Ser Tyr Val Asp Ala Leu Ile Gln Gln Gly
 145 150 155 160

Gln Gly Asp Asp Pro Glu Gly Leu Phe Ala Glu Ala Asn Ala Val Ala
 165 170 175

Cys Thr Leu Asp Met Val Met Ala Gly Thr Glu Thr Thr Ser Ala Thr
 180 185 190

Leu Gln Trp Ala Ala Leu Leu Met Gly Arg His Pro Asp Val Gln Gly
 195 200 205

Glu Thr Pro Ala Pro Gly Glu Thr Ala Pro Ser Ala Pro Gly Gly Pro
 210 215 220

Pro Gly Thr Arg Asp Gly Ala Ala Thr Gln Ala Ala Gln Pro Phe Ala
 225 230 235 240

Pro Gly Arg Val Gln Glu Glu Leu Asp Arg Val Leu Gly Pro Gly Arg
 245 250 255

Thr Pro Arg Leu Glu Asp Gln Gln Ala Leu Pro Tyr Thr Ser Ala Val
 260 265 270

Leu His Glu Val Gln Arg Phe Ile Thr Leu Leu Pro His Val Pro Arg
 275 280 285

Cys Thr Ala Ala Asp Thr Gln Leu Gly Gly Phe Leu Leu Pro Lys Gly
 290 295 300

Thr Pro Val Ile Pro Leu Leu Thr Ser Val Leu Leu Asp Glu Thr Gln
 305 310 315 320

Trp Gln Thr Pro Gly Gln Phe Asn Pro Gly His Phe Leu Asp Ala Asn
325 330 335

Gly His Phe Val Lys Arg Glu Ala Phe Leu Pro Phe Ser Ala Gly Arg
340 345 350

Arg Val Cys Val Gly Glu Arg Leu Ala Arg Thr Glu Leu Phe Leu Leu
355 360 365

Phe Ala Gly Leu Leu Gln Arg Tyr Arg Leu Leu Pro Pro Pro Gly Val
370 375 380

Ser Pro Ala Ser Leu Asp Thr Thr Pro Ala Arg Ala Phe Thr Met Arg
385 390 395 400

Pro Arg Ala Gln Ala Leu Cys Ala Val Pro Arg Pro
405 410

<210>	148
<211>	203
<212>	PRT
<213>	Homo sapien

<400> 148

Asp Pro Gly Ala Trp Arg Asp Gly Ser Phe Cys Pro Arg Gly Thr Pro
1 5 10 15

Arg Asp Glu Gly Trp Arg Cys His Pro Ser Gly Pro Pro Phe Ala Pro
20 25 30

Gly Arg Val Gln Glu Glu Leu Asp Arg Val Leu Gly Pro Gly Arg Thr
35 40 45

Pro Arg Leu Glu Asp Gln Gln Ala Leu Pro Tyr Thr Ser Ala Val Leu
50 55 60

His Glu Val Gln Arg Phe Ile Thr Leu Leu Pro His Val Pro Arg Cys
65 70 75 80

Thr Ala Ala Asp Thr Gln Leu Gly Gly Phe Leu Leu Pro Lys Gly Thr
85 90 95

Pro Val Ile Pro Leu Leu Thr Ser Val Leu Leu Asp Glu Thr Gln Trp
100 105 110

Gln Thr Pro Gly Gln Phe Asn Pro Gly His Phe Leu Asp Ala Asn Gly

208

115

120

125

His Phe Val Lys Arg Glu Ala Phe Leu Pro Phe Ser Ala Gly Arg Arg
 130 135 140

Val Cys Val Gly Glu Arg Leu Ala Arg Thr Glu Leu Phe Leu Leu Phe
 145 150 155 160

Ala Gly Leu Leu Gln Arg Tyr Arg Leu Leu Pro Pro Pro Gly Val Ser
 165 170 175

Pro Ala Ser Leu Asp Thr Thr Pro Ala Arg Ala Phe Thr Met Arg Pro
 180 185 190

Arg Ala Gln Ala Leu Cys Ala Val Pro Arg Pro
 195 200

<210> 149
 <211> 116
 <212> PRT
 <213> Homo sapien

<400> 149

Met Ala Arg Asp Ile Val Ala Met Ser Arg Ala Met Cys Leu Met Leu
 1 5 10 15

Leu Ser Val Ala Arg Ala Phe Leu Leu Met Val Val Arg Thr Glu Glu
 20 25 30

Val Ala Gly Phe Arg Trp Pro Asp Leu Arg Phe Asn Asp His His Asp
 35 40 45

Thr Phe Ala Val Gly Cys Arg Leu His Ala His Ser Leu Ala Val Asn
 50 55 60

Gln Ser Val Val Ala Glu Gly Ile Ala Gly Pro Gln Val Ile Gly Leu
 65 70 75 80

Ser Ala Val Val Phe Gly Leu Ser Phe Glu Asn Met Glu Asn Trp Ser
 85 90 95

Ser Ser Ala Arg Pro Ile Gln Leu Leu Met Pro Glu His Arg Tyr Ala
 100 105 110

Asp Ile Arg Gln
 115

209

<210> 150
 <211> 141
 <212> PRT
 <213> Homo sapien

<400> 150

Gly Glu Arg Pro Leu Ser Trp Ser Pro Leu Gly Arg Gly His Leu Cys
 1 5 10 15

Leu Val Pro Leu Gly Gly Arg Arg Gly Ala Cys Ala Gly Lys Ser Arg
 20 25 30

Arg Pro Arg Trp Ala Asp His Glu Val Arg Ser Ser Arg Pro Ala Trp
 35 40 45

Pro Thr Trp His Thr Trp His Ala His Arg Gly Asp Val Ala Cys His
 50 55 60

Val Ser Asp Ala Ala Glu Arg Gly Ala Ser Ile Leu Val Asp Gly Gly
 65 70 75 80

Pro His Ile Gly Gly Gly Arg Leu Pro Leu Ala Gly Ser Pro Leu Asn
 85 90 95

Asp His His Asp Thr Phe Ala Val Gly Cys Arg Leu His Ala His Ser
 100 105 110

Leu Ala Val Asn Gln Ser Val Val Ala Glu Gly Ile Ala Gly Pro Gln
 115 120 125

Val Ile Gly Leu Ser Ala Val Val Phe Gly Leu Ser Phe
 130 135 140

<210> 151
 <211> 426
 <212> PRT
 <213> Homo sapien

<400> 151

Met Ser Pro Ala Pro Asp Ala Ala Pro Ala Pro Ala Ser Ile Ser Leu
 1 5 10 15

Phe Asp Leu Ser Ala Asp Ala Pro Val Phe Gln Gly Leu Ser Leu Val
 20 25 30

Ser His Ala Pro Gly Glu Ala Leu Ala Arg Ala Pro Arg Thr Ser Cys
 35 40 45

210

Ser Gly Ser Gly Glu Arg Glu Ser Pro Glu Arg Lys Leu Leu Gln Gly
 50 55 60

Pro Met Asp Ile Ser Glu Lys Leu Phe Cys Ser Thr Cys Asp Gln Thr
 65 70 75 80

Phe Gln Asn His Gln Glu Gln Arg Glu His Tyr Lys Leu Asp Trp His
 85 90 95

Arg Phe Asn Leu Lys Gln Arg Leu Lys Asp Lys Pro Leu Leu Ser Ala
 100 105 110

Leu Asp Phe Glu Lys Gln Ser Ser Thr Gly Asp Leu Ser Ser Ile Ser
 115 120 125

Gly Ser Glu Asp Ser Asp Ser Ala Ser Glu Glu Asp Leu Gln Thr Leu
 130 135 140

Asp Arg Glu Arg Ala Thr Phe Glu Lys Leu Ser Arg Pro Pro Gly Phe
 145 150 155 160

Tyr Pro His Arg Val Leu Phe Gln Asn Ala Gln Gly Gln Phe Leu Tyr
 165 170 175

Ala Tyr Arg Cys Val Leu Gly Pro His Gln Asp Pro Pro Glu Glu Ala
 180 185 190

Glu Leu Leu Leu Gln Asn Leu Gln Ser Arg Gly Pro Arg Asp Cys Val
 195 200 205

Val Leu Met Ala Ala Ala Gly His Phe Ala Gly Ala Ile Phe Gln Gly
 210 215 220

Arg Glu Val Val Thr His Lys Thr Phe His Arg Tyr Thr Val Arg Ala
 225 230 235 240

Lys Arg Gly Thr Ala Gln Gly Leu Arg Asp Ala Arg Gly Gly Pro Ser
 245 250 255

His Ser Ala Gly Ala Asn Leu Arg Arg Tyr Asn Glu Ala Thr Leu Tyr
 260 265 270

Lys Asp Val Arg Asp Leu Leu Ala Gly Pro Ser Trp Ala Lys Ala Leu
 275 280 285

211

Glu Glu Ala Gly Thr Ile Leu Leu Arg Ala Pro Arg Ser Gly Arg Ser
 290 295 300

Leu Phe Phe Gly Gly Lys Gly Ala Pro Leu Gln Arg Gly Asp Pro Arg
 305 310 315 320

Leu Trp Asp Ile Pro Leu Ala Thr Arg Arg Pro Thr Phe Gln Glu Leu
 325 330 335

Gln Arg Val Leu His Lys Leu Thr Thr Leu His Val Tyr Glu Glu Asp
 340 345 350

Pro Arg Glu Ala Val Arg Leu His Ser Pro Gln Thr His Trp Lys Thr
 355 360 365

Val Arg Glu Glu Arg Lys Lys Pro Thr Glu Glu Glu Ile Arg Lys Ile
 370 375 380

Cys Arg Asp Glu Lys Glu Ala Leu Gly Gln Asn Glu Glu Ser Pro Lys
 385 390 395 400

Gln Gly Leu Ile Thr Ile Trp Gln Leu Ser Asp Leu Ser Phe Cys Pro
 405 410 415

Lys Asn Ala Leu Ala Asn Ser Leu Leu Ser
 420 425

<210> 152

<211> 370

<212> PRT

<213> Homo sapien

<400> 152

Met Ser Pro Ala Pro Asp Ala Ala Pro Ala Pro Ala Ser Ile Ser Leu
 1 5 10 15

Phe Asp Leu Ser Ala Asp Ala Pro Val Phe Gln Gly Leu Ser Leu Arg
 20 25 30

Glu His Tyr Lys Leu Asp Trp His Arg Phe Asn Leu Lys Gln Arg Leu
 35 40 45

Lys Asp Lys Pro Leu Leu Ser Ala Leu Asp Phe Glu Lys Gln Ser Ser
 50 55 60

Thr Gly Asp Leu Ser Ser Ile Ser Gly Ser Glu Asp Ser Asp Ser Ala

212

65	70	75	80
Ser Glu Glu Asp	Leu Gln Thr Leu Asp	Arg Glu Arg Ala Thr	Phe Glu
	85	90	95
Lys Leu Ser Arg	Pro Pro Gly Phe Tyr	Pro His Arg Val	Leu Phe Gln
	100	105	110
Asn Ala Gln Gly	Gln Phe Leu Tyr Ala Tyr	Arg Cys Val	Leu Gly Pro
	115	120	125
His Gln Asp Pro	Pro Glu Glu Ala Glu	Leu Leu Leu	Gln Asn Leu Gln
	130	135	140
Ser Arg Gly Pro	Arg Asp Cys Val Val	Leu Met Ala Ala Ala	Gly His
	145	150	155
Phe Ala Gly Ala	Ile Phe Gln Gly Arg	Glu Val Val Thr	His Lys Thr
	165	170	175
Phe His Arg Tyr	Thr Val Arg Ala Lys	Arg Gly Thr Ala	Gln Gly Leu
	180	185	190
Arg Asp Ala Arg	Gly Gly Pro Ser His	Ser Ala Gly Ala	Asn Leu Arg
	195	200	205
Arg Tyr Asn Glu	Ala Thr Leu Tyr Lys	Asp Val Arg Asp	Leu Leu Ala
	210	215	220
Gly Pro Ser Trp	Ala Lys Ala Leu Glu	Glu Ala Gly Thr	Ile Leu Leu
	225	230	235
Arg Ala Pro Arg	Ser Gly Arg Ser Leu	Phe Phe Gly Gly	Lys Gly Ala
	245	250	255
Pro Leu Gln Arg	Gly Asp Pro Arg Leu	Trp Asp Ile Pro	Leu Ala Thr
	260	265	270
Arg Arg Pro Thr	Phe Gln Glu Leu Gln	Arg Val Leu His	Lys Leu Thr
	275	280	285
Thr Leu His Val	Tyr Glu Glu Asp Pro	Arg Glu Ala Val	Arg Leu His
	290	295	300
Ser Pro Gln Thr	His Trp Lys Thr Val	Arg Glu Glu Arg	Lys Lys Pro
	305	310	315
			320

213

Thr Glu Glu Glu Ile Arg Lys Ile Cys Arg Asp Glu Lys Glu Ala Leu
 325 330 335

Gly Gln Asn Glu Glu Ser Pro Lys Gln Gly Leu Ile Thr Ile Trp Gln
 340 345 350

Leu Ser Asp Leu Ser Phe Cys Pro Lys Asn Ala Leu Ala Asn Ser Leu
 355 360 365

Leu Ser
 370

<210> 153
 <211> 208
 <212> PRT
 <213> Homo sapien

<400> 153

Met Ser Pro Ala Pro Asp Ala Ala Pro Ala Pro Ala Ser Ile Ser Leu
 1 5 10 15

Phe Asp Leu Ser Ala Asp Ala Pro Val Phe Gln Gly Leu Ser Leu Val
 20 25 30

Ser His Ala Pro Gly Glu Ala Leu Ala Arg Ala Pro Arg Thr Ser Cys
 35 40 45

Ser Gly Ser Gly Glu Arg Glu Ser Pro Glu Arg Lys Leu Leu Gln Gly
 50 55 60

Pro Met Asp Ile Ser Glu Lys Leu Phe Cys Ser Thr Cys Asp Gln Thr
 65 70 75 80

Phe Gln Asn His Gln Glu Gln Arg Glu His Tyr Lys Leu Asp Trp His
 85 90 95

Arg Phe Asn Leu Lys Gln Arg Leu Lys Asp Lys Pro Leu Leu Ser Ala
 100 105 110

Leu Asp Phe Glu Lys Gln Ser Ser Thr Gly Asp Leu Ser Ser Ile Ser
 115 120 125

Gly Ser Glu Asp Ser Asp Ser Ala Ser Glu Glu Asp Leu Gln Thr Leu
 130 135 140

214

Asp Arg Glu Arg Ala Thr Phe Glu Lys Leu Ser Arg Pro Pro Gly Phe
 145 150 155 160

Tyr Pro His Arg Val Leu Phe Gln Asn Ala Gln Gly Gln Phe Leu Tyr
 165 170 175

Ala Tyr Arg Cys Val Leu Gly Pro His Gln Arg Gln Val Thr Val Gln
 180 185 190

Val Ala Trp Leu Thr Pro Ala Phe Cys Thr Pro Ser Leu Asp Phe Pro
 195 200 205

<210> 154

<211> 209

<212> PRT

<213> Homo sapien

<400> 154

Trp Thr Gln Leu Leu Met Cys Tyr Phe Tyr Leu Gly Asp Lys Ile Lys
 1 5 10 15

Thr Ile Ser Phe Gln Ala Phe Ile Leu Met His Leu Leu Leu Pro Ser
 20 25 30

Glu Tyr Ser Leu Asp Gly Phe His Met Ser Gly Phe Ser Leu Gly Ser
 35 40 45

Gly Ser Glu Gly Glu Asp Gly Phe Gln Val Glu Leu Glu Leu Val Glu
 50 55 60

Leu Thr Val Gly Thr Leu Asp Leu Cys Glu Ser Glu Val Leu Pro Lys
 65 70 75 80

Arg Arg Arg Arg Lys Arg Asn Lys Lys Glu Lys Ser Arg Asp Gln Glu
 85 90 95

Ala Gly Ala His Arg Thr Leu Leu Gln Gln Thr Gln Glu Glu Glu Pro
 100 105 110

Ser Thr Gln Ser Ser Gln Ala Val Ala Ala Pro Leu Gly Pro Leu Leu
 115 120 125

Asp Glu Ala Lys Ala Pro Gly Gln Pro Glu Leu Trp Asn Ala Leu Leu
 130 135 140

Ala Ala Cys Arg Ala Gly Asp Val Gly Val Leu Lys Leu Gln Leu Ala
 145 150 155 160

215

Pro Ser Pro Ala Asp Pro Arg Val Leu Ser Leu Leu Ser Ala Pro Leu
 165 170 175

Gly Ser Gly Gly Phe Thr Leu Leu His Ala Ala Ala Ala Ala Gly Arg
 180 185 190

Gly Ser Val Val Arg Leu Leu Leu Glu Ala Gly Ala Asp Pro Thr Val
 195 200 205

Gln

<210> 155
 <211> 125
 <212> PRT
 <213> Homo sapien

<400> 155

Met Ser Pro Ala Pro Asp Ala Ala Pro Ala Pro Ala Ser Ile Ser Leu
 1 5 10 15

Phe Asp Leu Ser Ala Asp Ala Pro Val Phe Gln Gly Leu Ser Leu Val
 20 25 30

Ser His Ala Pro Gly Glu Ala Leu Ala Arg Ala Pro Arg Thr Ser Cys
 35 40 45

Ser Gly Ser Gly Glu Arg Glu Ser Pro Glu Arg Lys Leu Leu Gln Gly
 50 55 60

Pro Met Asp Ile Ser Glu Lys Leu Phe Cys Ser Thr Cys Asp Gln Thr
 65 70 75 80

Phe Gln Asn His Gln Glu Gln Arg Glu His Tyr Lys Leu Asp Trp His
 85 90 95

Arg Phe Asn Leu Lys Gln Arg Leu Lys Asp Lys Pro Leu Leu Ser Ala
 100 105 110

Leu Asp Phe Glu Lys Gln Ser Ser Thr Gly Asp Glu Trp
 115 120 125

<210> 156
 <211> 191
 <212> PRT
 <213> Homo sapien

216

<400> 156

Glu Pro Ser Leu Asp Arg Pro Gly Asp Asp Gln Leu Val Leu Gly Gly
 1 5 10 15

Gly Leu Cys Arg Val Glu Gly Ser Gln Val Pro Val Pro Ala Leu Ser
 20 25 30

Pro Ala Thr Ala Pro Thr Ser Phe Glu Gly Pro Phe Gly Lys Ile Val
 35 40 45

His Gln Val Arg Ala Ala Ile His Thr Pro Arg Phe Ser Lys Asp His
 50 55 60

Lys Cys Ser Leu Val Phe Tyr Ile Leu Ser Pro His Phe Leu Asp Pro
 65 70 75 80

Val Phe Leu Ser Thr Lys Ser His Ser Gln Arg Gln Pro Leu Leu Ala
 85 90 95

Thr Leu Ser Ser Val Pro Gly Ala Pro Glu Pro Cys Pro Gln Asp Gly
 100 105 110

Ser Pro Ala Ser His Pro Leu His Pro Pro Leu Cys Ile Ser Thr Gly
 115 120 125

Ala Thr Val Pro Tyr Phe Ala Glu Gly Ser Gly Gly Pro Val Pro Thr
 130 135 140

Thr Ser Thr Leu Ile Leu Pro Pro Glu Tyr Ser Ser Trp Gly Tyr Pro
 145 150 155 160

Tyr Gly Glu Ser Thr Ala Arg Ala Trp Gln Gly Gly Asp Ala Lys Ser
 165 170 175

Pro Thr Gln Thr Leu Leu Ser Ser Arg Arg Gly Pro Thr Val Leu
 180 185 190

<210> 157

<211> 130

<212> PRT

<213> Homo sapien

<400> 157

Met Gly Cys Leu Leu Thr Gly Leu Pro Arg Thr Leu Pro Arg Trp Cys
 1 5 10 15

217

Cys Leu Ala Pro Gly Arg Ile Pro Val Leu Ala Ala Ser Arg Gly Leu
 20 25 30

Gly Arg Arg Leu Ala Gly Ala His Ala Ala Ile Pro Phe Ala Ala Ile
 35 40 45

Arg Val Thr Cys Ile Gly Ser Cys Gly Val Ser Asn Lys Ala Asn Asp
 50 55 60

Thr Ala Trp Val Val Glu Glu Gly Tyr Phe Asn Ser Ser Leu Ser Leu
 65 70 75 80

Ala Asp Lys Gly Lys Phe Gly Ser Gln Phe Pro Ser Gly Asp Pro Trp
 85 90 95

Gly Gln Pro Leu Glu Trp Gly Leu Ser Val Leu Ser Ser Pro Phe Pro
 100 105 110

Arg Ser Ala Ser Trp Ile Trp His Trp Pro Ile Leu Ser Gln Gly Cys
 115 120 125

Gly Pro
 130

<210> 158
 <211> 340
 <212> PRT
 <213> Homo sapien

<400> 158

Pro Gly Glu Ala His Phe Arg Glu Asp His Trp Pro Ala Ala Gly Pro
 1 5 10 15

Thr Arg Arg Arg Ser Ser Arg Pro Gly Val Pro Leu Gln Gly Ala Glu
 20 25 30

Glu Asp Gly Ala Leu Trp Lys Gly Ala Arg Gly Phe Asn Gly Val Gln
 35 40 45

Leu Phe Glu Gly Met Lys Ala Phe Lys Gly Lys Asp Gln Gln Val Arg
 50 55 60

Leu Phe Arg Pro Trp Leu Asn Met Asp Arg Met Leu Arg Ser Ala Met
 65 70 75 80

Arg Leu Cys Leu Pro Ser Phe Asp Lys Leu Glu Leu Leu Glu Cys Ile

218

	85	90	95
Arg Arg Leu Ile Glu Val Asp Lys Asp Trp Val Pro Asp Ala Ala Gly	100	105	110
Thr Ser Leu Tyr Val Arg Pro Val Leu Ile Gly Asn Glu Pro Ser Leu	115	120	125
Gly Val Ser Gln Pro Thr Arg Ala Leu Leu Phe Val Ile Leu Cys Pro	130	135	140
Val Gly Ala Tyr Phe Pro Gly Gly Ser Val Thr Pro Val Ser Leu Leu	145	150	155
Ala Asp Pro Ala Phe Ile Arg Ala Trp Val Gly Gly Val Gly Asn Tyr	165	170	175
Lys Leu Gly Gly Asn Tyr Gly Pro Thr Val Leu Val Gln Gln Glu Ala	180	185	190
Leu Lys Arg Gly Cys Glu Gln Val Leu Trp Leu Tyr Gly Pro Asp His	195	200	205
Gln Leu Thr Glu Val Gly Thr Met Asn Ile Phe Val Tyr Trp Thr His	210	215	220
Glu Asp Gly Val Leu Glu Leu Val Thr Pro Pro Leu Asn Gly Val Ile	225	230	235
Leu Pro Gly Val Val Arg Gln Ser Leu Leu Asp Met Ala Gln Thr Trp	245	250	255
Gly Glu Phe Arg Val Val Glu Arg Thr Ile Thr Met Lys Gln Leu Leu	260	265	270
Arg Ala Leu Glu Glu Gly Arg Val Arg Glu Val Phe Gly Ser Gly Thr	275	280	285
Ala Cys Gln Val Cys Pro Val His Arg Ile Leu Tyr Lys Asp Arg Asn	290	295	300
Leu His Ile Pro Thr Met Glu Asn Gly Pro Glu Leu Ile Leu Arg Phe	305	310	315
Gln Lys Glu Leu Lys Glu Ile Gln Tyr Gly Ile Arg Ala His Glu Trp	325	330	335

219

Met Phe Pro Val
340

<210> 159
<211> 306
<212> PRT
<213> Homo sapien

<400> 159

Met Ala Pro Phe Gly Lys Glu His Glu Ala Leu Met Gly Glu Leu Phe
1 5 10 15

Glu Gly Met Lys Ala Phe Lys Gly Lys Asp Gln Gln Val Arg Leu Phe
20 25 30

Arg Pro Trp Leu Asn Met Asp Arg Met Leu Arg Ser Ala Met Arg Leu
35 40 45

Cys Leu Pro Ser Phe Asp Lys Leu Glu Leu Glu Cys Ile Arg Arg
50 55 60

Leu Ile Glu Val Asp Lys Asp Trp Val Pro Asp Ala Ala Gly Thr Ser
65 70 75 80

Leu Tyr Val Arg Pro Val Leu Ile Gly Asn Glu Pro Ser Leu Gly Val
85 90 95

Ser Gln Pro Thr Arg Ala Leu Leu Phe Val Ile Leu Cys Pro Val Gly
100 105 110

Ala Tyr Phe Pro Gly Gly Ser Val Thr Pro Val Ser Leu Leu Ala Asp
115 120 125

Pro Ala Phe Ile Arg Ala Trp Val Gly Gly Val Gly Asn Tyr Lys Leu
130 135 140

Gly Gly Asn Tyr Gly Pro Thr Val Leu Val Gln Gln Glu Ala Leu Lys
145 150 155 160

Arg Gly Cys Glu Gln Val Leu Trp Leu Tyr Gly Pro Asp His Gln Leu
165 170 175

Thr Glu Val Gly Thr Met Asn Ile Phe Val Tyr Trp Thr His Glu Asp
180 185 190

220

Gly Val Leu Glu Leu Val Thr Pro Pro Leu Asn Gly Val Ile Leu Pro
 195 200 205

Gly Val Val Arg Gln Ser Leu Leu Asp Met Ala Gln Thr Trp Gly Glu
 210 215 220

Phe Arg Val Val Glu Arg Thr Ile Thr Met Lys Gln Leu Leu Arg Ala
 225 230 235 240

Leu Glu Glu Gly Arg Val Arg Glu Val Phe Gly Ser Gly Thr Ala Cys
 245 250 255

Gln Val Cys Pro Val His Arg Ile Leu Tyr Lys Asp Arg Asn Leu His
 260 265 270

Ile Pro Thr Met Glu Asn Gly Pro Glu Leu Ile Leu Arg Phe Gln Lys
 275 280 285

Glu Leu Lys Glu Ile Gln Tyr Gly Ile Arg Ala His Glu Trp Met Phe
 290 295 300

Pro Val
 305

<210> 160
 <211> 485
 <212> PRT
 <213> Homo sapien

<400> 160

Gln Ile Val Tyr Leu Tyr Ile Gln Arg Ile Ile Arg Val Phe His Gly
 1 5 10 15

Val Asn Ala Pro Asp Asn Pro Leu Lys Glu Glu His Leu Val Gln Leu
 20 25 30

Asn Glu Thr Asp Ile Leu Arg Val Leu Asp Gly Asn Thr Gly Gly Thr
 35 40 45

Tyr Gly Gly His Ile Pro Gly Ser Asp Arg Ala Gly Leu Asn Arg His
 50 55 60

Asp Lys Ser Glu Asn Pro Gly Arg Ile Tyr Ala Gly Gly Ser Ser Thr
 65 70 75 80

Ala Ala Gly Asp Pro Ser Leu Leu Ser Ala Ala Arg Ile Met Ala Ala
 85 90 95

221

Ala Ala Leu Gly Gln Ile Trp Ala Arg Lys Leu Leu Ser Val Pro Trp
 100 105 110

Leu Leu Cys Gly Pro Arg Arg Tyr Ala Ser Ser Ser Phe Lys Ala Ala
 115 120 125

Asp Leu Gln Leu Glu Met Thr Gln Lys Pro His Lys Lys Pro Gly Pro
 130 135 140

Gly Glu Pro Leu Val Phe Gly Lys Thr Phe Thr Asp His Met Leu Met
 145 150 155 160

Val Glu Trp Asn Asp Lys Gly Trp Gly Gln Pro Arg Ile Gln Pro Phe
 165 170 175

Gln Asn Leu Thr Leu His Pro Ala Ser Ser Ser Leu His Tyr Ser Leu
 180 185 190

Gln Leu Phe Glu Gly Met Lys Ala Phe Lys Gly Lys Asp Gln Gln Val
 195 200 205

Arg Leu Phe Arg Pro Trp Leu Asn Met Asp Arg Met Leu Arg Ser Ala
 210 215 220

Met Arg Leu Cys Leu Pro Ser Phe Asp Lys Leu Glu Leu Leu Glu Cys
 225 230 235 240

Ile Arg Arg Leu Ile Glu Val Asp Lys Asp Trp Val Pro Asp Ala Ala
 245 250 255

Gly Thr Ser Leu Tyr Val Arg Pro Val Leu Ile Gly Asn Glu Pro Ser
 260 265 270

Leu Gly Val Ser Gln Pro Thr Arg Ala Leu Leu Phe Val Ile Leu Cys
 275 280 285

Pro Val Gly Ala Tyr Phe Pro Gly Gly Ser Val Thr Pro Val Ser Leu
 290 295 300

Leu Ala Asp Pro Ala Phe Ile Arg Ala Trp Val Gly Gly Val Gly Asn
 305 310 315 320

Tyr Lys Leu Gly Gly Asn Tyr Gly Pro Thr Val Leu Val Gln Gln Glu
 325 330 335

222

Ala Leu Lys Arg Gly Cys Glu Gln Val Leu Trp Leu Tyr Gly Pro Asp
 340 345 350

His Gln Leu Thr Glu Val Gly Thr Met Asn Ile Phe Val Tyr Trp Thr
 355 360 365

His Glu Asp Gly Val Leu Glu Leu Val Thr Pro Pro Leu Asn Gly Val
 370 375 380

Ile Leu Pro Gly Val Val Arg Gln Ser Leu Leu Asp Met Ala Gln Thr
 385 390 395 400

Trp Val Arg Thr Trp His Leu Leu Val Met Gly Ala Met Cys Gln Gly
 405 410 415

Pro Gly His Gln Arg Ala Gly Thr Gly Ala His Trp His Val Ser Ala
 420 425 430

Pro Ser Pro Gly Ser Val Ser Pro Val Gly Pro Leu Ser Phe Ser Leu
 435 440 445

Ser Ser Gly Arg Glu Arg Trp Arg Ser Ala Ala Gly Gln Pro Ser Gly
 450 455 460

Asp Thr Cys Leu Cys Gln Leu Pro Cys Arg Val Ser Ser Gly Trp Trp
 465 470 475 480

Ser Ala Arg Ser Pro
 485

<210> 161
 <211> 465
 <212> PRT
 <213> Homo sapien

<400> 161

Met Ala Ala Ala Ala Leu Gly Gln Ile Trp Ala Arg Lys Leu Leu Ser
 1 5 10 15

Val Pro Trp Leu Leu Cys Gly Pro Arg Arg Tyr Ala Ser Ser Ser Phe
 20 25 30

Lys Ala Ala Asp Leu Gln Leu Glu Met Thr Gln Lys Pro His Lys Lys
 35 40 45

Pro Gly Pro Gly Glu Pro Leu Val Phe Gly Lys Thr Phe Thr Asp His

223

50

55

60

Met Leu Met Val Glu Trp Asn Asp Lys Gly Trp Gly Gln Pro Arg Ile
 65 70 75 80

Gln Pro Phe Gln Asn Leu Thr Leu His Pro Ala Ser Ser Ser Leu His
 85 90 95

Tyr Ser Leu Gln Leu Phe Glu Gly Met Lys Ala Phe Lys Gly Lys Asp
 100 105 110

Gln Gln Val Arg Leu Phe Arg Pro Trp Leu Asn Met Asp Arg Met Leu
 115 120 125

Arg Ser Ala Met Arg Leu Cys Leu Pro Ser Phe Asp Lys Leu Glu Leu
 130 135 140

Leu Glu Cys Ile Arg Arg Leu Ile Glu Val Asp Lys Asp Trp Val Pro
 145 150 155 160

Asp Ala Ala Gly Thr Ser Leu Tyr Val Arg Pro Val Leu Ile Gly Asn
 165 170 175

Glu Pro Ser Leu Gly Val Ser Gln Pro Thr Arg Ala Leu Leu Phe Val
 180 185 190

Ile Leu Cys Pro Val Gly Ala Tyr Phe Pro Gly Gly Ser Val Thr Pro
 195 200 205

Val Ser Leu Leu Ala Asp Pro Ala Phe Ile Arg Ala Trp Val Gly Gly
 210 215 220

Val Gly Asn Tyr Lys Leu Gly Gly Asn Tyr Gly Pro Thr Val Leu Val
 225 230 235 240

Gln Gln Glu Ala Leu Lys Arg Gly Cys Glu Gln Val Leu Trp Leu Tyr
 245 250 255

Gly Pro Asp His Gln Leu Thr Glu Val Gly Thr Met Asn Ile Phe Val
 260 265 270

Tyr Trp Thr His Glu Asp Gly Val Leu Glu Leu Val Thr Pro Pro Leu
 275 280 285

Asn Gly Val Ile Leu Pro Gly Val Val Arg Gln Ser Leu Leu Asp Met
 290 295 300

224

Ala Gln Thr Trp Val Glu Asp Met Ala Ser Ser Gly Asp Gly Arg His
 305 310 315 320
 Val Pro Gly Ala Arg Ala Ser Glu Gly Trp Asp Trp Gly Thr Leu Ala
 325 330 335
 Cys Leu Cys Pro Phe Ser Trp Val Cys Leu Ser Arg Trp Ala Ser Val
 340 345 350
 Phe Leu Thr Ile Leu Arg Glu Gly Glu Val Glu Val Cys Ser Arg Ala
 355 360 365
 Thr Leu Trp Gly His Val Ser Leu Pro Thr Ala Leu Gln Gly Glu Phe
 370 375 380
 Arg Val Val Glu Arg Thr Ile Thr Met Lys Gln Leu Leu Arg Ala Leu
 385 390 395 400
 Glu Glu Gly Arg Val Arg Glu Val Phe Gly Ser Gly Thr Ala Cys Gln
 405 410 415
 Val Cys Pro Val His Arg Ile Leu Tyr Lys Asp Arg Asn Leu His Ile
 420 425 430
 Pro Thr Met Glu Asn Gly Pro Glu Leu Ile Leu Arg Phe Gln Lys Glu
 435 440 445
 Leu Lys Glu Ile Gln Tyr Gly Ile Arg Ala His Glu Trp Met Phe Pro
 450 455 460
 Val
 465

<210> 162
 <211> 74
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (3)..(5)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (8)..(10)
 <223> X=any amino acid

225

<220>
<221> MISC_FEATURE
<222> (13)..(14)
<223> X=any amino acid

<400> 162

Thr Leu Xaa Xaa Xaa Phe Asn Xaa Xaa Xaa Ser Ser Xaa Xaa Lys Ile
1 5 10 15

Arg Lys Asn Thr Ala Ser Tyr Val Pro Lys Glu Lys Lys Ile Lys Gly
20 25 30

Thr Met Pro Thr Cys Ser Thr Ile Lys Ala Ser Phe Ser Tyr Phe Phe
35 40 45

Asn Thr Lys Tyr Lys Gln Arg Ile His Ile Leu Lys Thr Glu Leu Arg
50 55 60

Ser Arg His Ala Val Leu Glu Thr Leu Gln
65 70

<210> 163
<211> 63
<212> PRT
<213> Homo sapien

<400> 163

Thr Arg Val Leu Gln Cys Ala Pro Arg Cys Ser Pro Asn Cys Val Ala
1 5 10 15

Phe Thr Ala Ala Val Thr Val Pro Ala Cys Ile Tyr Ala Leu Phe Gly
20 25 30

Pro Cys Glu His Thr Gly Ile Leu Val Ile Leu Pro Pro Met Glu Tyr
35 40 45

Leu Trp Arg Ser Pro Val Phe Ile Tyr Phe Gly Ile Asn Pro Leu
50 55 60

<210> 164
<211> 313
<212> PRT
<213> Homo sapien

<400> 164

Met Lys Cys Glu His Cys Thr Arg Lys Glu Cys Ser Lys Lys Thr Lys

226

1	5	10	15
Thr	Asp	Asp	Gln
20		25	30
Thr	Asp	Asp	Gln
20		25	30
Glu	Asn	Gly	Glu
35		40	45
Glu	Asn	Gly	Glu
35		40	45
Phe	Leu	Ser	Asp
50		55	60
Phe	Leu	Ser	Asp
50		55	60
Gly	Val	Gln	Leu
65		70	75
Gly	Val	Gln	Leu
65		70	75
Asn	Leu	Asn	Lys
85		90	95
Asn	Leu	Asn	Lys
85		90	95
Ser	Val	Cys	Pro
100		105	110
Ser	Val	Cys	Pro
100		105	110
Ser	Val	Thr	Asp
115		120	125
Ser	Val	Thr	Asp
115		120	125
Gly	Val	His	Tyr
130		135	140
Gly	Val	His	Tyr
130		135	140
Leu	Glu	Ser	Gln
145		150	155
Leu	Glu	Ser	Gln
145		150	155
Glu	Glu	Arg	Thr
165		170	175
Glu	Glu	Arg	Thr
165		170	175
Arg	Tyr	Ala	Glu
180		185	190
Arg	Tyr	Ala	Glu
180		185	190
Thr	Ala	Lys	Ser
195		200	205
Thr	Ala	Lys	Ser
195		200	205
Phe	Ala	Arg	Gln
210		215	220
Phe	Ala	Arg	Gln
210		215	220
Val	Ala	Pro	Cys
225		230	235
Val	Ala	Pro	Cys
225		230	235
Pro	Pro	Ala	Leu
245		250	255
Pro	Pro	Ala	Leu
245		250	255

227

Gly Glu Ile Ala Gln Ile Met Glu Gln Gly Asp Leu Ser Val Arg Asp
 260 265 270

Ala Ala Val Asp Thr Leu Val Ser Gly Phe Ser Gly Glu Ser Pro Leu
 275 280 285

Gly Gly Arg Thr Ser Arg Gln Pro Cys Gln Pro Pro Ala Arg Pro Arg
 290 295 300

Ala Ala Leu Leu Tyr Asp Glu Ala Met
 305 310

<210> 165
 <211> 395
 <212> PRT
 <213> Homo sapien

<400> 165

Glu Pro Arg Val Arg Arg Val Ser Asn Ala Glu Leu Ala Asp Arg Ala
 1 5 10 15

Arg Pro Arg Pro Pro Arg Ala Gln Gly Pro Pro Gly Pro Val Thr Thr
 20 25 30

Gly Pro Ser Thr Leu Glu Arg Pro Gln Leu Gly Leu Gly Thr Val Arg
 35 40 45

Ala Leu Thr Asp Ser Leu Val Asn Ala Ala Trp Pro Pro Pro Pro Pro
 50 55 60

Gln Asp Pro Arg Glu Ala Glu Thr Gly Ala Arg Thr Arg Ser Pro Arg
 65 70 75 80

Arg Arg Thr Trp Ser Glu Pro Ala Ala Pro Pro Arg Ala Leu Arg Leu
 85 90 95

Ala Leu Gly Pro Gly Pro Pro Leu Pro Asp Thr Val Ile Gly Leu Gly
 100 105 110

Lys Ala Val Arg Val Gly Asn Pro Ile Gly Pro Gly Val Arg Leu Val
 115 120 125

Leu Ser Arg Cys Ser His Trp Pro Ser Ala Ala Val Gly Gly Glu Ala
 130 135 140

228

Ala Ser Gly Glu Asp Asn Lys Gly Pro Arg Ala Ala Gly Ser Gly Val
 145 150 155 160

Pro Val Ser Arg Cys Phe Pro Glu Ala Glu Ala Pro Gly Leu Pro Pro
 165 170 175

Ala Ala Leu Gln Met Lys Cys Glu His Cys Thr Arg Lys Glu Cys Ser
 180 185 190

Lys Lys Thr Lys Thr Asp Asp Gln Glu Asn Val Ser Ala Asp Ala Pro
 195 200 205

Ser Pro Ala Gln Glu Asn Gly Glu Lys Gly Glu Phe His Lys Leu Ala
 210 215 220

Asp Ala Lys Ile Phe Leu Ser Asp Cys Leu Ala Cys Asp Ser Cys Met
 225 230 235 240

Thr Ala Glu Glu Gly Val Gln Leu Ser Gln Gln Asn Ala Lys Asp Phe
 245 250 255

Phe Arg Val Leu Asn Leu Asn Lys Lys Cys Asp Thr Ser Lys His Lys
 260 265 270

Val Leu Val Val Ser Val Cys Pro Gln Ser Leu Pro Tyr Phe Ala Ala
 275 280 285

Lys Phe Asn Leu Ser Val Thr Asp Ala Ser Arg Arg Leu Cys Gly Phe
 290 295 300

Leu Lys Ser Leu Gly Val His Tyr Val Phe Asp Thr Thr Ile Ala Ala
 305 310 315 320

Asp Phe Ser Ile Leu Glu Ser Gln Lys Glu Phe Val Arg Arg Tyr Arg
 325 330 335

Gln His Ser Glu Glu Glu Arg Thr Leu Pro Met Leu Thr Ser Ala Cys
 340 345 350

Pro Gly Trp Val Arg Tyr Ala Glu Arg Val Leu Gly Arg Pro Ile Thr
 355 360 365

Ala His Leu Cys Thr Ala Lys Ser Pro Gln Gln Val Met Gly Ser Leu
 370 375 380

Val Lys Asp Tyr Phe Ala Arg Gln Gln Val Ser

229

385

390

395

<210> 166
 <211> 285
 <212> PRT
 <213> Homo sapien

<400> 166

Met Gly Ser Ala Phe Pro Val Leu Pro Gly Gly Ser Thr Gly Trp Gly
 1 5 10 15

Ala Leu Gln Met Phe Gly Arg Thr Thr Pro Ser Pro Glu Gly Gly Ser
 20 25 30

Arg Gln Thr Trp Met Glu Cys Trp Cys Ser Asn Leu Ser Pro Glu Lys
 35 40 45

Ile Phe His Val Ile Val Ala Pro Cys Tyr Asp Lys Lys Leu Glu Ala
 50 55 60

Leu Gln Glu Ser Leu Pro Pro Ala Leu His Gly Ser Arg Gly Ala Asp
 65 70 75 80

Cys Val Leu Thr Ser Gly Glu Ile Ala Gln Ile Met Glu Gln Gly Asp
 85 90 95

Leu Ser Val Arg Asp Ala Ala Val Asp Thr Leu Phe Gly Asp Leu Lys
 100 105 110

Glu Asp Lys Val Thr Arg His Asp Gly Ala Ser Ser Asp Gly His Leu
 115 120 125

Ala His Ile Phe Arg His Ala Ala Lys Glu Leu Phe Asn Glu Asp Val
 130 135 140

Glu Glu Val Thr Tyr Arg Ala Leu Arg Asn Lys Asp Phe Gln Glu Val
 145 150 155 160

Thr Leu Glu Lys Asn Gly Glu Val Val Leu Arg Phe Ala Ala Ala Tyr
 165 170 175

Gly Phe Arg Asn Ile Gln Asn Met Ile Leu Lys Leu Lys Lys Gly Lys
 180 185 190

Phe Pro Phe His Phe Val Glu Val Leu Ala Cys Ala Gly Gly Cys Leu
 195 200 205

230

Asn Gly Arg Gly Gln Ala Gln Thr Pro Asp Gly His Ala Asp Lys Ala
 210 215 220

Leu Leu Arg Gln Met Glu Gly Ile Tyr Ala Asp Ile Pro Val Arg Arg
 225 230 235 240

Pro Glu Ser Ser Ala His Val Gln Glu Leu Tyr Gln Glu Trp Leu Glu
 245 250 255

Gly Ile Asn Ser Pro Lys Ala Arg Glu Val Leu His Thr Thr Tyr Gln
 260 265 270

Ser Gln Glu Arg Gly Thr His Ser Leu Asp Ile Lys Trp
 275 280 285

<210> 167
 <211> 170
 <212> PRT
 <213> Homo sapien

<400> 167

Asp Ser Val Ser His Pro Ala Lys Lys Phe Ala Met Ser Ala Ala Lys
 1 5 10 15

Cys Lys Gln Ser Arg Leu Trp Val Lys Ala Leu Val Met Lys Asn Lys
 20 25 30

Lys Lys Lys Ile Pro Lys His Phe Leu Thr Leu Val Leu Lys Lys Thr
 35 40 45

Pro Ser Pro Lys Val Cys Phe Ser Phe Asn Phe Leu Asp Phe Ala Arg
 50 55 60

Leu Ser Gly Cys Lys Leu Pro Thr Leu Phe Trp Met Asp Arg Asp Lys
 65 70 75 80

Val Phe Lys Gln Arg Leu Cys Pro Leu His Lys Pro Phe Pro Pro Pro
 85 90 95

Pro Pro Gln Pro Pro Ala Ala Ile Ile Thr Gly Ala Val Lys Trp Leu
 100 105 110

Leu Ser Asp Gly His Thr Thr Arg Arg Gln Met Lys Arg Ser Gly Ser
 115 120 125

Lys Arg Gly Ser Ala Ser Gln Tyr Gln Pro Ala Val Pro Arg Gly Gly

231

130

135

140

Ser Ala Gly Arg Thr Val Phe Pro Arg Gln Ala Ala Met Pro Pro Pro
 145 150 155 160

Thr Ala Lys Ala Pro Lys Ala Thr Ser Val
 165 170

<210> 168
 <211> 159
 <212> PRT
 <213> Homo sapien

<400> 168

Met Ser Ala Ala Lys Cys Lys Gln Ser Arg Leu Trp Val Lys Ala Leu
 1 5 10 15

Val Met Lys Asn Lys Lys Lys Lys Ile Pro Lys His Phe Leu Thr Leu
 20 25 30

Val Leu Lys Lys Thr Pro Ser Pro Lys Val Cys Phe Ser Phe Asn Phe
 35 40 45

Leu Asp Phe Ala Arg Leu Ser Gly Cys Lys Leu Pro Thr Leu Phe Trp
 50 55 60

Met Asp Arg Asp Lys Val Phe Lys Gln Arg Leu Cys Pro Leu His Lys
 65 70 75 80

Pro Phe Pro Pro Pro Pro Pro Gln Pro Pro Ala Ala Ile Ile Thr Gly
 85 90 95

Ala Val Lys Trp Leu Leu Ser Asp Gly His Thr Thr Arg Arg Gln Met
 100 105 110

Lys Arg Ser Gly Ser Lys Arg Gly Ser Ala Ser Gln Tyr Gln Pro Ala
 115 120 125

Val Pro Arg Gly Gly Ser Ala Gly Arg Thr Val Phe Pro Arg Gln Ala
 130 135 140

Ala Met Pro Pro Pro Thr Ala Lys Ala Pro Lys Gly Asn Ile Arg
 145 150 155

<210> 169
 <211> 170
 <212> PRT

232

<213> Homo sapien

<400> 169

Asp Ser Val Ser His Pro Ala Lys Lys Phe Ala Met Ser Ala Ala Lys
 1 5 10 15

Cys Lys Gln Ser Arg Leu Trp Val Lys Ala Leu Val Met Lys Asn Lys
 20 25 30

Lys Lys Lys Ile Pro Lys His Phe Leu Thr Leu Val Leu Lys Lys Thr
 35 40 45

Pro Ser Pro Lys Val Cys Phe Ser Phe Asn Phe Leu Asp Phe Ala Arg
 50 55 60

Leu Ser Gly Cys Lys Leu Pro Thr Leu Phe Trp Met Asp Arg Asp Lys
 65 70 75 80

Val Phe Lys Gln Arg Leu Cys Pro Leu His Lys Pro Phe Pro Pro Pro
 85 90 95

Pro Pro Gln Pro Pro Ala Ala Ile Ile Thr Gly Ala Val Lys Trp Leu
 100 105 110

Leu Ser Asp Gly His Thr Thr Arg Arg Gln Met Lys Arg Ser Gly Ser
 115 120 125

Lys Arg Gly Ser Ala Ser Gln Tyr Gln Pro Ala Val Pro Arg Gly Gly
 130 135 140

Ser Ala Gly Arg Thr Val Phe Pro Arg Gln Ala Ala Met Pro Pro Pro
 145 150 155 160

Thr Ala Lys Ala Pro Lys Ala Thr Ser Val
 165 170

<210> 170

<211> 255

<212> PRT

<213> Homo sapien

<400> 170

Gln Leu Leu Arg Asp Pro Asn Val Ala Leu Glu Leu Ser Ala Met Cys
 1 5 10 15

Ser Thr Val Pro Trp Arg Arg Thr Leu Arg Glu Gly Gln Pro Cys His
 20 25 30

233

Leu Ser Leu Pro His His His Ser Pro Pro Pro Ile Lys Leu Gln Ser
 35 40 45

Gly Cys Trp Thr Pro Leu Gly Ala Val Ser Ala His His Pro Leu Cys
 50 55 60

Ala Ala Thr Trp Ser Gln Ala His Cys Pro Leu Ala Gly Arg Gly Pro
 65 70 75 80

Ser Arg Arg Arg Cys Gly Leu His Arg Ala Pro Ser Thr Lys Glu Ser
 85 90 95

Ala Asn Ala Ser Ala Gly Pro Arg Ala Met Ala Ser Leu Pro Gln Leu
 100 105 110

Met Ala Ala Pro Thr Ser Ser Cys Thr Ser Leu Met Trp Lys Gly Ser
 115 120 125

Met Ser Gln Trp Lys Ala Thr Arg Ser Pro Ile Lys Cys Ala Pro Ser
 130 135 140

His Pro Arg Met Arg Ser Cys Arg Pro Trp Arg Ser Ser Ser Leu Thr
 145 150 155 160

Trp His Gln Ala Pro Ser Met Arg Pro Gly Leu Asp Met Ser Ser Ala
 165 170 175

Pro Arg Arg Trp Trp Lys His Pro Leu Ser Cys Ala Cys Gly Arg Leu
 180 185 190

Cys Gly Glu Glu Ala Ala Asp Thr Gly Asp Asp Ile Leu Pro His Glu
 195 200 205

Thr Gly Leu Gln Pro Gly Met Val Pro Leu Lys Tyr Leu Leu Glu Glu
 210 215 220

Gly Val Trp Gly Ala Gly Val Gly Cys Gly Val Phe Pro Ala Ile Ser
 225 230 235 240

Thr Ala Tyr Asp His Cys Asn Asn Leu Ser Pro Ser Glu Glu His
 245 250 255

<210> 171
 <211> 147
 <212> PRT

234

<213> Homo sapien

<400> 171

Met Ser Ser Glu Pro Pro Pro Pro Pro Gln Pro Pro Thr His Gln Ala
1 5 10 15

Ser Val Gly Leu Leu Asp Thr Pro Arg Ser Arg Glu Arg Ser Pro Ser
20 25 30

Pro Leu Arg Gly Asn Val Val Pro Ser Pro Leu Pro Thr Arg Arg Thr
35 40 45

Arg Thr Phe Ser Ala Thr Val Arg Ala Ser Gln Gly Pro Val Tyr Lys
50 55 60

Gly Val Cys Lys Cys Phe Cys Arg Ser Lys Gly His Gly Phe Ile Thr
65 70 75 80

Pro Ala Asp Gly Gly Pro Asp Ile Phe Leu His Ile Ser Asp Val Glu
85 90 95

Gly Glu Tyr Val Pro Val Glu Gly Asp Glu Val Thr Tyr Lys Met Cys
100 105 110

Ser Ile Pro Pro Lys Asn Glu Lys Leu Gln Ala Val Glu Val Val Ile
115 120 125

Thr His Leu Ala Pro Gly Thr Lys His Glu Thr Trp Ser Gly His Val
130 135 140

Ile Ser Ser
145

<210> 172

<211> 255

<212> PRT

<213> Homo sapien

<400> 172

Gln Leu Leu Arg Asp Pro Asn Val Ala Leu Glu Leu Ser Ala Met Cys
1 5 10 15

Ser Thr Val Pro Trp Arg Arg Thr Leu Arg Glu Gly Gln Pro Cys His
20 25 30

Leu Ser Leu Pro His His His Ser Pro Pro Pro Ile Lys Leu Gln Ser
35 40 45

235

Gly Cys Trp Thr Pro Leu Gly Ala Val Ser Ala His His Pro Leu Cys
 50 55 60

Ala Ala Thr Trp Ser Gln Ala His Cys Pro Leu Ala Gly Arg Gly Pro
 65 70 75 80

Ser Arg Arg Arg Cys Gly Leu His Arg Ala Pro Ser Thr Lys Glu Ser
 85 90 95

Ala Asn Ala Ser Ala Gly Pro Arg Ala Met Ala Ser Leu Pro Gln Leu
 100 105 110

Met Ala Ala Pro Thr Ser Ser Cys Thr Ser Leu Met Trp Lys Gly Ser
 115 120 125

Met Ser Gln Trp Lys Ala Thr Arg Ser Pro Ile Lys Cys Ala Pro Ser
 130 135 140

His Pro Arg Met Arg Ser Cys Arg Pro Trp Arg Ser Ser Ser Leu Thr
 145 150 155 160

Trp His Gln Ala Pro Ser Met Arg Pro Gly Leu Asp Met Ser Ser Ala
 165 170 175

Pro Arg Arg Trp Trp Lys His Pro Leu Ser Cys Ala Cys Gly Arg Leu
 180 185 190

Cys Gly Glu Glu Ala Ala Asp Thr Gly Asp Asp Ile Leu Pro His Glu
 195 200 205

Thr Gly Leu Gln Pro Gly Met Val Pro Leu Lys Tyr Leu Leu Glu Glu
 210 215 220

Gly Val Trp Gly Ala Gly Val Gly Cys Gly Val Phe Pro Ala Ile Ser
 225 230 235 240

Thr Ala Tyr Asp His Cys Asn Asn Leu Ser Pro Ser Glu Glu His
 245 250 255

<210> 173

<211> 243

<212> PRT

<213> Homo sapien

<400> 173

236

Leu Arg Thr Gly Arg Asn Ser Gly Gly Gly Gln Asn Gly Leu Gln Gly
 1 5 10 15
 Gln Pro Cys His Leu Ser Leu Pro His His His Ser Pro Pro Pro Ile
 20 25 30
 Lys Leu Gln Ser Gly Cys Trp Thr Pro Leu Gly Ala Val Ser Ala His
 35 40 45
 His Pro Leu Cys Ala Ala Thr Trp Ser Gln Ala His Cys Pro Leu Ala
 50 55 60
 Gly Arg Gly Pro Ser Arg Arg Arg Cys Gly Leu His Arg Ala Pro Ser
 65 70 75 80
 Thr Lys Glu Ser Ala Asn Ala Ser Ala Gly Pro Arg Ala Met Ala Ser
 85 90 95
 Leu Pro Gln Leu Met Ala Ala Pro Thr Ser Ser Cys Thr Ser Leu Met
 100 105 110
 Trp Lys Gly Ser Met Ser Gln Trp Lys Ala Thr Arg Ser Pro Ile Lys
 115 120 125
 Cys Ala Pro Ser His Pro Arg Met Arg Ser Cys Arg Pro Trp Arg Ser
 130 135 140
 Ser Ser Leu Thr Trp His Gln Ala Pro Ser Met Arg Pro Gly Leu Asp
 145 150 155 160
 Met Ser Ser Ala Pro Arg Arg Trp Trp Lys His Pro Leu Ser Cys Ala
 165 170 175
 Cys Gly Arg Leu Cys Gly Glu Glu Ala Ala Asp Thr Gly Asp Asp Ile
 180 185 190
 Leu Pro His Glu Thr Gly Leu Gln Pro Gly Met Val Pro Leu Lys Tyr
 195 200 205
 Leu Leu Glu Glu Gly Val Trp Gly Ala Gly Val Gly Cys Gly Val Phe
 210 215 220
 Pro Ala Ile Ser Thr Ala Tyr Asp His Cys Asn Asn Leu Ser Pro Ser
 225 230 235 240
 Glu Glu His

237

<210> 174
 <211> 147
 <212> PRT
 <213> Homo sapien

<400> 174

Met Ser Ser Glu Pro Pro Pro Pro Pro Gln Pro Pro Thr His Gln Ala
 1 5 10 15

Ser Val Gly Leu Leu Asp Thr Pro Arg Ser Arg Glu Arg Ser Pro Ser
 20 25 30

Pro Leu Arg Gly Asn Val Val Pro Ser Pro Leu Pro Thr Arg Arg Thr
 35 40 45

Arg Thr Phe Ser Ala Thr Val Arg Ala Ser Gln Gly Pro Val Tyr Lys
 50 55 60

Gly Val Cys Lys Cys Phe Cys Arg Ser Lys Gly His Gly Phe Ile Thr
 65 70 75 80

Pro Ala Asp Gly Gly Pro Asp Ile Phe Leu His Ile Ser Asp Val Glu
 85 90 95

Gly Glu Tyr Val Pro Val Glu Gly Asp Glu Val Thr Tyr Lys Met Cys
 100 105 110

Ser Ile Pro Pro Lys Asn Glu Lys Leu Gln Ala Val Glu Val Val Ile
 115 120 125

Thr His Leu Ala Pro Gly Thr Lys His Glu Thr Trp Ser Gly His Val
 130 135 140

Ile Ser Ser
 145

<210> 175
 <211> 202
 <212> PRT
 <213> Homo sapien

<400> 175

Trp Gly Thr Ala Val Gly Arg Gly Trp Asn Glu Leu Cys Val Pro Arg
 1 5 10 15

238

Ser Ala Asp Gly Ala Pro Asp Ile Ser Pro Ser Leu Ser Gly Arg Cys
 20 25 30

Gly Leu His Arg Ala Pro Ser Thr Lys Glu Ser Ala Asn Ala Ser Ala
 35 40 45

Gly Pro Arg Ala Met Ala Ser Leu Pro Gln Leu Met Ala Ala Pro Thr
 50 55 60

Ser Ser Cys Thr Ser Leu Met Trp Lys Gly Ser Met Ser Gln Trp Lys
 65 70 75 80

Ala Thr Arg Ser Pro Ile Lys Cys Ala Pro Ser His Pro Arg Met Arg
 85 90 95

Ser Cys Arg Pro Trp Arg Ser Ser Ser Leu Thr Trp His Gln Ala Pro
 100 105 110

Ser Met Arg Pro Gly Leu Asp Met Ser Ser Ala Pro Arg Arg Trp Trp
 115 120 125

Lys His Pro Leu Ser Cys Ala Cys Gly Arg Leu Cys Gly Glu Glu Ala
 130 135 140

Ala Asp Thr Gly Asp Asp Ile Leu Pro His Glu Thr Gly Leu Gln Pro
 145 150 155 160

Gly Met Val Pro Leu Lys Tyr Leu Leu Glu Glu Gly Val Trp Gly Ala
 165 170 175

Gly Val Gly Cys Gly Val Phe Pro Ala Ile Ser Thr Ala Tyr Asp His
 180 185 190

Cys Asn Asn Leu Ser Pro Ser Glu Glu His
 195 200

<210> 176
 <211> 138
 <212> PRT
 <213> Homo sapien

<400> 176

Met Ala Ser Leu Pro Gln Leu Met Ala Ala Pro Thr Ser Ser Cys Thr
 1 5 10 15

Ser Leu Met Trp Lys Gly Ser Met Ser Gln Trp Lys Ala Thr Arg Ser
 20 25 30

239

Pro Ile Lys Cys Ala Pro Ser His Pro Arg Met Arg Ser Cys Arg Pro
 35 40 45

Trp Arg Ser Ser Ser Leu Thr Trp His Gln Ala Pro Ser Met Arg Pro
 50 55 60

Gly Leu Asp Met Ser Ser Ala Pro Arg Arg Trp Trp Lys His Pro Leu
 65 70 75 80

Ser Cys Ala Cys Gly Arg Leu Cys Gly Glu Glu Ala Ala Asp Thr Gly
 85 90 95

Asp Asp Ile Leu Pro His Glu Thr Gly Leu Gln Pro Gly Met Val Pro
 100 105 110

Leu Lys Tyr Leu Leu Glu Glu Gly Val Trp Gly Gly Arg Cys Gly Val
 115 120 125

Trp Gly Val Pro Gly His Gln His Ser Leu
 130 135

<210> 177

<211> 185

<212> PRT

<213> Homo sapien

<400> 177

Met Ser Ser Glu Pro Pro Pro Pro Gln Pro Pro Thr His Gln Ala
 1 5 10 15

Ser Val Gly Leu Leu Asp Thr Pro Arg Ser Arg Glu Arg Ser Pro Ser
 20 25 30

Pro Ser Thr Lys Glu Ser Ala Asn Ala Ser Ala Gly Pro Arg Ala Met
 35 40 45

Ala Ser Leu Pro Gln Leu Met Ala Ala Pro Thr Ser Ser Cys Thr Ser
 50 55 60

Leu Met Trp Lys Gly Ser Met Ser Gln Trp Lys Ala Thr Arg Ser Pro
 65 70 75 80

Ile Lys Cys Ala Pro Ser His Pro Arg Met Arg Ser Cys Arg Pro Trp
 85 90 95

240

Arg Ser Ser Ser Leu Thr Trp His Gln Ala Pro Ser Met Arg Pro Gly
 100 105 110

Leu Asp Met Ser Ser Ala Pro Arg Arg Trp Trp Lys His Pro Leu Ser
 115 120 125

Cys Ala Cys Gly Arg Leu Cys Gly Glu Glu Ala Ala Asp Thr Gly Asp
 130 135 140

Asp Ile Leu Pro His Glu Thr Gly Leu Gln Pro Gly Met Val Pro Leu
 145 150 155 160

Lys Tyr Leu Leu Glu Glu Gly Val Trp Gly Gly Arg Cys Gly Val Trp
 165 170 175

Gly Val Pro Gly His Gln His Ser Leu
 180 185

<210> 178
 <211> 265
 <212> PRT
 <213> Homo sapien

<400> 178

Ser Phe Pro Pro Ala His Leu Phe Ser Ala Cys Arg Gly Ser Ser Ser
 1 5 10 15

Arg Pro Pro Arg Cys Phe Cys Leu Trp Ala Gly Ala Leu Asp Gly Gly
 20 25 30

Leu Ala Gly Arg Trp Gly Glu Ala Arg Gly Ala Ser His Ala Gly Ser
 35 40 45

Arg Ala Thr Pro Arg Arg Ala Trp Pro Arg Gln Leu Trp Leu Glu Val
 50 55 60

Gly Thr Ser Ala Met Ser Ser Glu Pro Pro Pro Pro Gln Pro Pro
 65 70 75 80

Thr His Gln Ala Ser Val Gly Leu Leu Asp Thr Pro Arg Ser Arg Glu
 85 90 95

Arg Ser Pro Ser Pro Ser Thr Lys Glu Ser Ala Asn Ala Ser Ala Gly
 100 105 110

Pro Arg Ala Met Ala Ser Leu Pro Gln Leu Met Ala Ala Pro Thr Ser
 115 120 125

241

Ser Cys Thr Ser Leu Met Trp Lys Gly Ser Met Ser Gln Trp Lys Ala
 130 135 140

Thr Arg Ser Pro Ile Lys Cys Ala Pro Ser His Pro Arg Met Arg Ser
 145 150 155 160

Cys Arg Pro Trp Arg Ser Ser Ser Leu Thr Trp His Gln Ala Pro Ser
 165 170 175

Met Arg Pro Gly Leu Asp Met Ser Ser Ala Pro Arg Arg Trp Trp Lys
 180 185 190

His Pro Leu Ser Cys Ala Cys Gly Arg Leu Cys Gly Glu Glu Ala Ala
 195 200 205

Asp Thr Gly Asp Asp Ile Leu Pro His Glu Thr Gly Leu Gln Pro Gly
 210 215 220

Met Val Pro Leu Lys Tyr Leu Leu Glu Glu Gly Val Trp Gly Ala Gly
 225 230 235 240

Val Gly Cys Gly Val Phe Pro Ala Ile Ser Thr Ala Tyr Asp His Cys
 245 250 255

Asn Asn Leu Ser Pro Ser Glu Glu His
 260 265

<210> 179
 <211> 201
 <212> PRT
 <213> Homo sapien

<400> 179

Met Cys Ser Thr Val Pro Trp Arg Arg Thr Leu Arg Glu Gly Gln Phe
 1 5 10 15

Leu Pro Leu Leu Pro Cys Gly Val Trp Leu Pro Ala Ala Ser Gly Arg
 20 25 30

Val Arg Gly Val Ala Glu Phe Gly Ser Arg Trp Leu Ala Leu Lys Ser
 35 40 45

Pro Trp Leu Trp Val Phe Phe Phe Glu Thr Glu Ser Cys Ser Val Ala
 50 55 60

242

Gln Ala Gly Val Gln Trp Cys Asp Leu Ser Ser Leu Glu Pro Pro Pro
65 70 75 80

Pro Arg Phe Lys Gln Phe Ser Cys Leu Ser Leu Gln Val Asp Gly Ile
85 90 95

Thr Gly Ala Cys His His Ala Gln Leu Ile Phe Val Phe Val Leu Glu
100 105 110

Thr Gly Phe Pro His Val Gly Gln Ala Ser Leu Glu Leu Leu Thr Leu
115 120 125

Ser Asp Pro Pro Ala Ser Ala Ser Gln Ser Ala Gly Ile Ala Gly Val
130 135 140

Ser His Cys Ala Arg Pro Leu Ala Leu Gly Phe Leu Leu Thr Phe Leu
145 150 155 160

Leu Pro Ser His Lys Tyr Tyr Val Ser Gln Met Cys Arg Asp Pro Cys
165 170 175

Leu Val Leu Gly Thr Gln Arg Gly Pro Met Pro Ser Trp Ser Trp Gly
180 185 190

Arg Met Trp His Phe His Glu Glu Leu
195 200

<210> 180
<211> 159
<212> PRT
<213> Homo sapien

<400> 180

Met Ala Gln Gly Leu Ala Val Arg Glu Met Thr Gly Met Thr Lys Phe
1 5 10 15

Lys Pro Tyr Phe Ile Ser Ala Thr Ser Glu Ile Leu Ser Gln Lys Cys
20 25 30

Ile Asn Thr Asn Val Leu Phe Leu Ser Leu Ser Asp Asn His Gly His
35 40 45

Ile Asp Pro Ser Leu Arg Leu Ile Trp Asp Leu Ala Phe Leu Gly Ser
50 55 60

Ser Tyr Val Met Trp Glu Met Thr Thr Gln Val Ser His Tyr Tyr Leu
65 70 75 80

243

Ala Gln Leu Thr Ser Val Arg Gln Trp Lys Thr Asn Asp Asp Thr Ile
 85 90 95

Asp Phe Asp Tyr Thr Val Leu Leu His Glu Leu Ser Thr Gln Glu Ile
 100 105 110

Ile Pro Cys Arg Ile His Leu Val Trp Tyr Pro Gly Lys Pro Leu Lys
 115 120 125

Val Lys Tyr His Cys Gln Glu Leu Gln Thr Pro Glu Glu Ala Ser Gly
 130 135 140

Thr Glu Glu Gly Ser Ala Val Val Pro Thr Glu Leu Ser Asn Phe
 145 150 155

<210> 181
 <211> 128
 <212> PRT
 <213> Homo sapien

<400> 181

Met Arg Ile His Asp Phe His Ile Leu Lys Gln Asn Thr Thr Lys Asn
 1 5 10 15

Arg Glu Ala Glu Ile Glu Lys Ala Val Gly Asp Thr Arg His Pro Phe
 20 25 30

Lys Thr Arg Met Tyr Cys Ile Val Ile Thr Leu Asn Thr Thr Ile Phe
 35 40 45

Ile Thr Leu Thr Leu Phe Ser Pro Ser Arg Lys Thr Asn Asp Asp Thr
 50 55 60

Ile Asp Phe Asp Tyr Thr Val Leu Leu His Glu Leu Ser Thr Gln Glu
 65 70 75 80

Ile Ile Pro Cys Arg Ile His Leu Val Trp Tyr Pro Gly Lys Pro Leu
 85 90 95

Lys Val Lys Tyr His Cys Gln Glu Leu Gln Thr Pro Glu Glu Ala Ser
 100 105 110

Gly Thr Glu Glu Gly Ser Ala Val Val Pro Thr Glu Leu Ser Asn Phe
 115 120 125

244

<210> 182
 <211> 224
 <212> PRT
 <213> Homo sapien

<400> 182

Met Asp Gly Asp Gln Thr Glu Glu Arg Met Met Lys Met Met Val His
 1 5 10 15

Gln Arg Pro Leu Pro Gln Pro Ala Leu Leu Pro Met Ser Ser Asn Thr
 20 25 30

Phe Pro Ser Arg Ser Thr Lys Pro Ser Pro Met Asn Pro Leu Pro Ser
 35 40 45

Ser His Met Pro Gly Ala Phe Ser Glu Ser Asn Ser Ser Phe Pro Gln
 50 55 60

Ser Ala Ser Leu Pro Pro Tyr Phe Ser Gln Gly Pro Ser Asn Arg Pro
 65 70 75 80

Pro Ile Arg Ala Glu Gly Arg Asn Phe Pro Leu Pro Leu Pro Asn Lys
 85 90 95

Pro Arg Pro Pro Ser Pro Ala Glu Glu Glu Asn Ser Leu Asn Glu Glu
 100 105 110

Trp Tyr Val Ser Tyr Ile Thr Arg Pro Glu Ala Glu Ala Ala Leu Arg
 115 120 125

Lys Ile Asn Gln Asp Gly Thr Phe Leu Val Arg Asp Ser Ser Lys Lys
 130 135 140

Thr Thr Thr Asn Pro Tyr Val Leu Met Val Leu Tyr Lys Asp Lys Val
 145 150 155 160

Tyr Asn Ile Gln Ile Arg Tyr Gln Lys Glu Ser Gln Val Tyr Leu Leu
 165 170 175

Gly Thr Gly Leu Arg Gly Lys Glu Asp Phe Leu Ser Val Ser Asp Ile
 180 185 190

Ile Asp Tyr Phe Arg Lys Met Pro Leu Leu Leu Ile Asp Gly Lys Asn
 195 200 205

Arg Gly Ser Arg Tyr Gln Cys Thr Leu Thr His Ala Ala Gly Tyr Pro
 210 215 220

245

<210> 183
 <211> 230
 <212> PRT
 <213> Homo sapien

<400> 183

Met Phe Leu Ser Asp Ser Trp Gly Leu Pro Ile Ser Phe Ile Val Phe
 1 5 10 15

Tyr Phe Cys Thr Val His Gln Arg Pro Leu Pro Gln Pro Ala Leu Leu
 20 25 30

Pro Met Ser Ser Asn Thr Phe Pro Ser Arg Ser Thr Lys Pro Ser Pro
 35 40 45

Met Asn Pro Leu Pro Ser Ser His Met Pro Gly Ala Phe Ser Glu Ser
 50 55 60

Asn Ser Ser Phe Pro Gln Ser Ala Ser Leu Pro Pro Tyr Phe Ser Gln
 65 70 75 80

Gly Pro Ser Asn Arg Pro Pro Ile Arg Ala Glu Gly Arg Asn Phe Pro
 85 90 95

Leu Pro Leu Pro Asn Lys Pro Arg Pro Pro Ser Pro Ala Glu Glu Glu
 100 105 110

Asn Ser Leu Asn Glu Glu Trp Tyr Val Ser Tyr Ile Thr Arg Pro Glu
 115 120 125

Ala Glu Ala Ala Leu Arg Lys Ile Asn Gln Asp Gly Thr Phe Leu Val
 130 135 140

Arg Asp Ser Ser Lys Lys Thr Thr Thr Asn Pro Tyr Val Leu Met Val
 145 150 155 160

Leu Tyr Lys Asp Lys Val Tyr Asn Ile Gln Ile Arg Tyr Gln Lys Glu
 165 170 175

Ser Gln Val Tyr Leu Leu Gly Thr Gly Leu Arg Gly Lys Glu Asp Phe
 180 185 190

Leu Ser Val Ser Asp Ile Ile Asp Tyr Phe Arg Lys Met Pro Leu Leu
 195 200 205

246

Leu Ile Asp Gly Lys Asn Arg Gly Ser Arg Tyr Gln Cys Thr Leu Thr
 210 215 220

His Ala Ala Gly Tyr Pro
 225 230

<210> 184
 <211> 225
 <212> PRT
 <213> Homo sapien

<400> 184

Gln Leu Gly Pro Ala Asn Leu Ile His Cys Phe Leu Leu Leu His Val
 1 5 10 15

His Gln Arg Pro Leu Pro Gln Pro Ala Leu Leu Pro Met Ser Ser Asn
 20 25 30

Thr Phe Pro Ser Arg Ser Thr Lys Pro Ser Pro Met Asn Pro Leu Pro
 35 40 45

Ser Ser His Met Pro Gly Ala Phe Ser Glu Ser Asn Ser Ser Phe Pro
 50 55 60

Gln Ser Ala Ser Leu Pro Pro Tyr Phe Ser Gln Gly Pro Ser Asn Arg
 65 70 75 80

Pro Pro Ile Arg Ala Glu Gly Arg Asn Phe Pro Leu Pro Leu Pro Asn
 85 90 95

Lys Pro Arg Pro Pro Ser Pro Ala Glu Glu Glu Asn Ser Leu Asn Glu
 100 105 110

Glu Trp Tyr Val Ser Tyr Ile Thr Arg Pro Glu Ala Glu Ala Ala Leu
 115 120 125

Arg Lys Ile Asn Gln Asp Gly Thr Phe Leu Val Arg Asp Ser Ser Lys
 130 135 140

Lys Thr Thr Thr Asn Pro Tyr Val Leu Met Val Leu Tyr Lys Asp Lys
 145 150 155 160

Val Tyr Asn Ile Gln Ile Arg Tyr Gln Lys Glu Ser Gln Val Tyr Leu
 165 170 175

Leu Gly Thr Gly Leu Arg Gly Lys Glu Asp Phe Leu Ser Val Ser Asp
 180 185 190

247

Ile Ile Asp Tyr Phe Arg Lys Met Pro Leu Leu Leu Ile Asp Gly Lys
 195 200 205

Asn Arg Gly Ser Arg Tyr Gln Cys Thr Leu Thr His Ala Ala Gly Tyr
 210 215 220

Pro
 225

<210> 185
 <211> 1085
 <212> PRT
 <213> Homo sapien

<400> 185

Met Ala Ala Ser Thr Gly Tyr Val Arg Leu Trp Gly Ala Ala Arg Cys
 1 5 10 15

Trp Val Leu Arg Arg Pro Met Leu Ala Ala Ala Gly Gly Arg Val Pro
 20 25 30

Thr Ala Ala Gly Ala Trp Leu Leu Arg Gly Gln Arg Thr Cys Asp Ala
 35 40 45

Ser Pro Pro Trp Ala Leu Trp Gly Arg Gly Pro Ala Ile Gly Gly Gln
 50 55 60

Trp Arg Gly Phe Trp Glu Ala Ser Ser Arg Gly Gly Gly Ala Phe Ser
 65 70 75 80

Gly Gly Glu Asp Ala Ser Glu Gly Gly Ala Glu Glu Gly Ala Gly Gly
 85 90 95

Ala Gly Gly Ser Ala Gly Ala Gly Glu Gly Pro Val Ile Thr Ala Leu
 100 105 110

Thr Pro Met Thr Ile Pro Asp Val Phe Pro His Leu Pro Leu Ile Ala
 115 120 125

Ile Thr Arg Asn Pro Val Phe Pro Arg Phe Ile Lys Ile Ile Glu Val
 130 135 140

Lys Asn Lys Lys Leu Val Glu Leu Leu Arg Arg Lys Val Arg Leu Ala
 145 150 155 160

248

Gln Pro Tyr Val Gly Val Phe Leu Lys Arg Asp Asp Ser Asn Glu Ser
 165 170 175

Asp Val Val Glu Ser Leu Asp Glu Ile Tyr His Thr Gly Thr Phe Ala
 180 185 190

Gln Ile His Glu Met Gln Asp Leu Gly Asp Lys Leu Arg Met Ile Val
 195 200 205

Met Gly His Arg Arg Val His Ile Ser Arg Gln Leu Glu Val Glu Pro
 210 215 220

Glu Glu Pro Glu Ala Glu Asn Lys His Lys Pro Arg Arg Lys Ser Lys
 225 230 235 240

Arg Gly Lys Lys Glu Ala Glu Asp Glu Leu Ser Ala Arg His Pro Ala
 245 250 255

Glu Leu Ala Met Glu Pro Thr Pro Glu Leu Pro Ala Glu Val Leu Met
 260 265 270

Val Glu Val Glu Asn Val Val His Glu Asp Phe Gln Val Thr Glu Glu
 275 280 285

Val Lys Ala Leu Thr Ala Glu Ile Val Lys Thr Ile Arg Asp Ile Ile
 290 295 300

Ala Leu Asn Pro Leu Tyr Arg Glu Ser Val Leu Gln Met Met Gln Ala
 305 310 315 320

Gly Gln Arg Val Val Asp Asn Pro Ile Tyr Leu Ser Asp Met Gly Ala
 325 330 335

Ala Leu Thr Gly Ala Glu Ser His Glu Leu Gln Asp Val Leu Glu Glu
 340 345 350

Thr Asn Ile Pro Lys Arg Leu Tyr Lys Ala Leu Ser Leu Leu Lys Lys
 355 360 365

Glu Phe Glu Leu Ser Lys Leu Gln Gln Arg Leu Gly Arg Glu Val Glu
 370 375 380

Glu Lys Ile Lys Gln Thr His Arg Lys Tyr Leu Leu Gln Glu Gln Leu
 385 390 395 400

Lys Ile Ile Lys Lys Glu Leu Gly Leu Glu Lys Asp Asp Lys Asp Ala

249

405 410 415

Ile Glu Glu Lys Phe Arg Glu Arg Leu Lys Glu Leu Val Val Pro Lys
 420 425 430

His Val Met Asp Val Val Asp Glu Glu Leu Ser Lys Leu Gly Leu Leu
 435 440 445

Asp Asn His Ser Ser Glu Phe Asn Val Thr Arg Asn Tyr Leu Asp Trp
 450 455 460

Leu Thr Ser Ile Pro Trp Gly Lys Tyr Ser Asn Glu Asn Leu Asp Leu
 465 470 475 480

Ala Arg Ala Gln Ala Val Leu Glu Glu Asp His Tyr Gly Met Glu Asp
 485 490 495

Val Lys Lys Arg Ile Leu Glu Phe Ile Ala Val Ser Gln Leu Arg Gly
 500 505 510

Ser Thr Gln Gly Lys Ile Leu Cys Phe Tyr Gly Pro Pro Gly Val Gly
 515 520 525

Lys Thr Ser Ile Ala Arg Ser Ile Ala Arg Ala Leu Asn Arg Glu Tyr
 530 535 540

Phe Arg Phe Ser Val Gly Gly Met Thr Asp Val Ala Glu Ile Lys Gly
 545 550 555 560

His Arg Arg Thr Tyr Val Gly Ala Met Pro Gly Lys Ile Ile Gln Cys
 565 570 575

Leu Lys Lys Thr Lys Thr Glu Asn Pro Leu Ile Leu Ile Asp Glu Val
 580 585 590

Arg Ala Glu Arg Pro Gly Asp Pro Leu Thr Arg Gln Cys Arg Val Gln
 595 600 605

Asp Glu Ala Glu Ala Phe Arg Val Leu Gly Ser Pro Ser Gln Leu Arg
 610 615 620

Gly His Arg Arg Asp Ala Gly Pro Asp Gly Thr Asp Gly Ser Leu Pro
 625 630 635 640

Thr His Ser His Ala Gln Gln Ala Ala Arg Pro Arg Gln Gly Val Gly
 645 650 655

250

Ala Ala Leu Thr Arg Gly Leu Pro Ser Pro Asp Pro Gly Val Ser Pro
 660 665 670

Leu Gly Gly Leu Ala Arg Gly Thr Ala Arg Gly Thr Thr Cys Cys Leu
 675 680 685

Leu Pro Gln Val Asp Lys Ile Gly Arg Gly Tyr Gln Gly Asp Pro Ser
 690 695 700

Ser Ala Leu Leu Glu Leu Leu Asp Pro Glu Gln Asn Ala Asn Phe Leu
 705 710 715 720

Asp His Tyr Leu Asp Val Pro Val Asp Leu Ser Lys Val Gly Gly Leu
 725 730 735

Ser Gly Ala Trp Ala Cys Trp Gly Gly Tyr Ala Ala Ser Pro Pro Ala
 740 745 750

Pro Cys Arg Arg Pro Gln Val Leu Phe Ile Cys Thr Ala Asn Val Thr
 755 760 765

Asp Thr Ile Pro Glu Pro Leu Arg Asp Arg Met Glu Met Ile Asn Val
 770 775 780

Ser Gly Tyr Val Ala Gln Glu Lys Leu Ala Ile Ala Glu Arg Tyr Leu
 785 790 795 800

Val Pro Gln Ala Arg Ala Leu Cys Gly Leu Asp Glu Ser Lys Ala Lys
 805 810 815

Leu Ser Ser Asp Val Leu Thr Leu Leu Ile Lys Gln Tyr Cys Arg Glu
 820 825 830

Ser Gly Val Arg Asn Leu Gln Lys Gln Val Glu Lys Val Leu Arg Lys
 835 840 845

Ser Ala Tyr Lys Ile Val Ser Gly Glu Ala Glu Ser Val Glu Val Thr
 850 855 860

Pro Glu Asn Leu Gln Asp Phe Val Gly Lys Pro Val Phe Thr Val Glu
 865 870 875 880

Arg Met Tyr Asp Val Thr Pro Pro Gly Val Val Met Gly Leu Ala Trp
 885 890 895

251

Thr Ala Met Gly Gly Ser Thr Leu Phe Val Glu Thr Ser Leu Arg Arg
 900 905 910

Pro Gln Asp Lys Asp Ala Lys Gly Asp Lys Asp Gly Ser Leu Glu Val
 915 920 925

Thr Gly Gln Leu Gly Glu Val Met Lys Glu Ser Ala Arg Ile Ala Tyr
 930 935 940

Thr Phe Ala Arg Ala Phe Leu Met Gln His Ala Pro Ala Asn Asp Tyr
 945 950 955 960

Leu Val Thr Ser His Ile His Leu His Val Pro Glu Gly Ala Thr Pro
 965 970 975

Lys Asp Gly Pro Ser Ala Gly Cys Thr Ile Val Thr Ala Leu Leu Ser
 980 985 990

Leu Ala Met Gly Arg Pro Val Arg Gln Asn Leu Ala Met Thr Gly Glu
 995 1000 1005

Val Ser Leu Thr Gly Lys Ile Leu Pro Val Gly Gly Ile Lys Glu
 1010 1015 1020

Lys Thr Ile Ala Ala Lys Arg Ala Gly Val Thr Cys Ile Val Leu
 1025 1030 1035

Pro Ala Glu Asn Lys Lys Asp Phe Tyr Asp Leu Ala Ala Phe Ile
 1040 1045 1050

Thr Glu Gly Leu Glu Val His Phe Val Glu His Tyr Arg Glu Ile
 1055 1060 1065

Phe Asp Ile Ala Phe Pro Asp Glu Gln Ala Glu Ala Leu Ala Val
 1070 1075 1080

Glu Arg
 1085

<210> 186

<211> 668

<212> PRT

<213> Homo sapien

<400> 186

Asn Ser Ala Pro Ser Ser Pro Arg Arg Pro Ser Ser Leu Lys Arg Leu

252

1	5	10	15
Phe	Arg	Ala	Thr
20	Arg	Pro	Ser
Gly	Thr	Glu	Ala
25	Arg	Ala	Gly
30	Arg	His	
Val	Arg	Phe	Ala
35	Ala	Ser	Gly
40	Asn	Asp	Ala
45	Ser	Cys	Val
50	Ser	Arg	Gln
55	Tyr	Gly	Tyr
60	Val	Arg	Leu
65	Trp	Gly	
70	Ala	Ala	Ala
75	Met	Leu	Ala
80	Ala	Gly	
85	Gln	Arg	
90	Arg	Gly	Gln
95	Arg	Gly	
100	Pro	Pro	Trp
105	Ala	Leu	Trp
110	Gly	Arg	Gly
115	Pro	Ala	Gly
120	Arg	Gly	
125	Ser	Arg	Gly
130	Gly	Glu	Gly
135	Ala	Glu	Gly
140	Gly	Glu	Gly
145	Pro	Val	
150	Gly	Glu	Gly
155	Gly	Glu	Gly
160	Val		
165	Leu	Thr	Pro
170	Met	Thr	Ile
175	Asp	Val	Phe
180	Pro	His	Leu
185	Pro	Arg	Phe
190	Ile	Lys	
195	Ile	Glu	Val
200	Lys	Asn	Lys
205	Leu	Val	Glu
210	Leu	Arg	Arg
215	Lys	Arg	Asp
220	Asp	Asp	
225	Ser	Asn	Glu
230	Ser	Asp	Val
235	Val	Glu	Ser
240	Leu	Asp	Glu
245	Ile	Tyr	His
250	Thr		
255	Lys	Leu	

253

Arg Met Ile Val Met Gly His Arg Arg Val His Ile Ser Arg Gln Leu
 260 265 270

Glu Val Glu Pro Glu Glu Pro Glu Ala Glu Asn Lys His Lys Pro Arg
 275 280 285

Arg Lys Ser Lys Arg Gly Lys Lys Glu Ala Glu Asp Glu Leu Ser Ala
 290 295 300

Arg His Pro Ala Glu Leu Ala Met Glu Pro Thr Pro Glu Leu Pro Ala
 305 310 315 320

Glu Val Leu Met Val Glu Val Glu Asn Val Val His Glu Asp Phe Gln
 325 330 335

Val Thr Glu Glu Val Lys Ala Leu Thr Ala Glu Ile Val Lys Thr Ile
 340 345 350

Arg Asp Ile Ile Ala Leu Asn Pro Leu Tyr Arg Glu Ser Val Leu Gln
 355 360 365

Met Met Gln Ala Gly Gln Arg Val Val Asp Asn Pro Ile Tyr Leu Ser
 370 375 380

Asp Met Gly Ala Ala Leu Thr Gly Ala Glu Ser His Glu Leu Gln Asp
 385 390 395 400

Val Leu Glu Glu Thr Asn Ile Pro Lys Arg Leu Tyr Lys Ala Leu Ser
 405 410 415

Leu Leu Lys Lys Glu Phe Glu Leu Ser Lys Leu Gln Gln Arg Leu Gly
 420 425 430

Arg Glu Val Glu Glu Lys Ile Lys Gln Thr His Arg Lys Tyr Leu Leu
 435 440 445

Gln Glu Gln Leu Lys Ile Ile Lys Lys Glu Leu Gly Leu Glu Lys Asp
 450 455 460

Asp Lys Asp Ala Ile Glu Glu Lys Phe Arg Glu Arg Leu Lys Glu Leu
 465 470 475 480

Val Val Pro Lys His Val Met Asp Val Val Asp Glu Glu Leu Ser Lys
 485 490 495

254

Leu Gly Leu Leu Asp Asn His Ser Ser Glu Phe Asn Val Thr Arg Asn
 500 505 510

Tyr Leu Asp Trp Leu Thr Ser Ile Pro Trp Gly Lys Tyr Ser Asn Glu
 515 520 525

Asn Leu Asp Leu Ala Arg Ala Gln Ala Val Leu Glu Glu Asp His Tyr
 530 535 540

Gly Met Glu Asp Val Lys Lys Arg Ile Leu Glu Phe Ile Ala Val Ser
 545 550 555 560

Gln Leu Arg Gly Ser Thr Gln Gly Lys Ile Leu Cys Phe Tyr Gly Pro
 565 570 575

Pro Gly Val Gly Lys Thr Ser Ile Ala Arg Ser Ile Ala Arg Ala Leu
 580 585 590

Asn Arg Glu Tyr Phe Arg Phe Ser Val Gly Gly Met Thr Asp Val Ala
 595 600 605

Glu Ile Lys Gly His Arg Arg Thr Tyr Val Gly Ala Met Pro Gly Lys
 610 615 620

Ile Ile Gln Cys Leu Lys Lys Thr Lys Thr Glu Asn Pro Leu Ile Leu
 625 630 635 640

Ile Asp Glu Val Arg Ala Glu Arg Pro Gly Asp Pro Leu Thr Arg Gln
 645 650 655

Cys Arg Val Gln Asp Glu Ala Glu Pro Ser Gly Ser
 660 665

<210> 187

<211> 771

<212> PRT

<213> Homo sapien

<400> 187

Met Glu Glu Arg Lys His Ser Leu Phe Phe Ser Pro Glu Ile Ala Tyr
 1 5 10 15

Pro Lys Asp Lys Ala Leu Arg Gly Lys Glu Thr Gly Phe Gly Phe Ser
 20 25 30

Phe Asp Ser Leu Gly Pro Val Ser Leu Gly Leu Val Ser Pro Pro Gly

255

35

40

45

Thr Ala Leu Leu His Leu Pro Gly Arg Pro Trp Pro Gly Pro Val Gly
 50 55 60

Gly Gly Glu Gly Val Asp Gly Ala Gly Gly Trp Gly Ser Gln Gly His
 65 70 75 80

Ala Pro Thr Gln Pro Thr Pro Cys Trp Ser Thr Ala Gly Pro Gln Arg
 85 90 95

Pro Cys Pro Ile Leu Gly Ala Leu Thr Ala Glu Ile Val Lys Thr Ile
 100 105 110

Arg Asp Ile Ile Ala Leu Asn Pro Leu Tyr Arg Glu Ser Val Leu Gln
 115 120 125

Met Met Gln Ala Gly Gln Arg Val Val Asp Asn Pro Ile Tyr Leu Ser
 130 135 140

Asp Met Gly Ala Ala Leu Thr Gly Ala Glu Ser His Glu Leu Gln Asp
 145 150 155 160

Val Leu Glu Glu Thr Asn Ile Pro Lys Arg Leu Tyr Lys Ala Leu Ser
 165 170 175

Leu Leu Lys Lys Glu Phe Glu Leu Ser Lys Leu Gln Gln Arg Leu Gly
 180 185 190

Arg Glu Val Glu Glu Lys Ile Lys Gln Thr His Arg Lys Tyr Leu Leu
 195 200 205

Gln Glu Gln Leu Lys Ile Ile Lys Lys Glu Leu Gly Leu Glu Lys Asp
 210 215 220

Asp Lys Asp Ala Ile Glu Glu Lys Phe Arg Glu Arg Leu Lys Glu Leu
 225 230 235 240

Val Val Pro Lys His Val Met Asp Val Val Asp Glu Glu Leu Ser Lys
 245 250 255

Leu Gly Leu Leu Asp Asn His Ser Ser Glu Phe Asn Val Thr Arg Asn
 260 265 270

Tyr Leu Asp Trp Leu Thr Ser Ile Pro Trp Gly Lys Tyr Ser Asn Glu
 275 280 285

256

Asn Leu Asp Leu Ala Arg Ala Gln Ala Val Leu Glu Glu Asp His Tyr
 290 295 300

Gly Met Glu Asp Val Lys Lys Arg Ile Leu Glu Phe Ile Ala Val Ser
 305 310 315 320

Gln Leu Arg Gly Ser Thr Gln Gly Lys Ile Leu Cys Phe Tyr Gly Pro
 325 330 335

Pro Gly Val Gly Lys Thr Ser Ile Ala Arg Ser Ile Ala Arg Ala Leu
 340 345 350

Asn Arg Glu Tyr Phe Arg Phe Ser Val Gly Gly Met Thr Asp Val Ala
 355 360 365

Glu Ile Lys Gly His Arg Arg Thr Tyr Val Gly Ala Met Pro Gly Lys
 370 375 380

Ile Ile Gln Cys Leu Lys Lys Thr Lys Thr Glu Asn Pro Leu Ile Leu
 385 390 395 400

Ile Asp Glu Val Asp Lys Ile Gly Arg Gly Tyr Gln Gly Asp Pro Ser
 405 410 415

Ser Ala Leu Leu Glu Leu Leu Asp Pro Glu Gln Asn Ala Asn Phe Leu
 420 425 430

Asp His Tyr Leu Asp Val Pro Val Asp Leu Ser Lys Val Leu Phe Ile
 435 440 445

Cys Thr Ala Asn Val Thr Asp Thr Ile Pro Glu Pro Leu Arg Asp Arg
 450 455 460

Met Glu Met Ile Asn Val Ser Gly Tyr Val Ala Gln Glu Lys Leu Ala
 465 470 475 480

Ile Ala Glu Arg Tyr Leu Val Pro Gln Ala Arg Ala Leu Cys Gly Leu
 485 490 495

Asp Glu Ser Lys Ala Lys Leu Ser Ser Asp Val Leu Thr Leu Leu Ile
 500 505 510

Lys Gln Tyr Cys Arg Glu Ser Gly Val Arg Asn Leu Gln Lys Gln Val
 515 520 525

257

Glu Lys Val Leu Arg Lys Ser Ala Tyr Lys Ile Val Ser Gly Glu Ala
 530 535 540

Glu Ser Val Glu Val Thr Pro Glu Asn Leu Gln Asp Phe Val Gly Lys
 545 550 555 560

Pro Val Phe Thr Val Glu Arg Met Tyr Asp Val Thr Pro Pro Gly Val
 565 570 575

Val Met Gly Leu Ala Trp Thr Ala Met Gly Gly Ser Thr Leu Phe Val
 580 585 590

Glu Thr Ser Leu Arg Arg Pro Gln Asp Lys Asp Ala Lys Gly Asp Lys
 595 600 605

Asp Gly Ser Leu Glu Val Thr Gly Gln Leu Gly Glu Val Met Lys Glu
 610 615 620

Ser Ala Arg Ile Ala Tyr Thr Phe Ala Arg Ala Phe Leu Met Gln His
 625 630 635 640

Ala Pro Ala Asn Asp Tyr Leu Val Thr Ser His Ile His Leu His Val
 645 650 655

Pro Glu Gly Ala Thr Pro Lys Asp Gly Pro Ser Ala Gly Cys Thr Ile
 660 665 670

Val Thr Ala Leu Leu Ser Leu Ala Met Gly Arg Pro Val Arg Gln Asn
 675 680 685

Leu Ala Met Thr Gly Glu Val Ser Leu Thr Gly Lys Ile Leu Pro Val
 690 695 700

Gly Gly Ile Lys Glu Lys Thr Ile Ala Ala Lys Arg Ala Gly Val Thr
 705 710 715 720

Cys Ile Val Leu Pro Ala Glu Asn Lys Lys Asp Phe Tyr Asp Leu Ala
 725 730 735

Ala Phe Ile Thr Glu Gly Leu Glu Val His Phe Val Glu His Tyr Arg
 740 745 750

Glu Ile Phe Asp Ile Ala Phe Pro Asp Glu Gln Ala Glu Ala Leu Ala
 755 760 765

258

Val Glu Arg
770

<210> 188
<211> 848
<212> PRT
<213> Homo sapien

<400> 188

Met Ala Ala Ser Thr Gly Tyr Val Arg Leu Trp Gly Ala Ala Arg Cys
1 5 10 15

Trp Val Leu Arg Arg Pro Met Leu Ala Ala Gly Gly Arg Val Pro
20 25 30

Thr Ala Ala Gly Ala Trp Leu Leu Arg Gly Gln Arg Thr Cys Asp Ala
35 40 45

Ser Pro Pro Trp Ala Leu Trp Gly Arg Gly Pro Ala Ile Gly Gly Gln
50 55 60

Trp Arg Gly Phe Trp Glu Ala Ser Ser Arg Gly Gly Gly Ala Phe Ser
65 70 75 80

Gly Gly Glu Asp Ala Ser Glu Gly Gly Ala Glu Glu Gly Ala Gly Gly
85 90 95

Ala Gly Gly Ser Ala Gly Ala Gly Glu Gly Pro Val Ile Thr Ala Leu
100 105 110

Thr Pro Met Thr Ile Pro Asp Val Phe Pro His Leu Pro Leu Ile Ala
115 120 125

Ile Thr Arg Asn Pro Val Phe Pro Arg Phe Ile Lys Ile Ile Glu Val
130 135 140

Lys Asn Lys Lys Leu Val Glu Leu Leu Arg Arg Lys Val Arg Leu Ala
145 150 155 160

Gln Pro Tyr Val Gly Val Phe Leu Lys Arg Asp Asp Ser Asn Glu Ser
165 170 175

Asp Val Val Glu Ser Leu Asp Glu Ile Tyr His Thr Gly Thr Phe Ala
180 185 190

Gln Ile His Glu Met Gln Asp Leu Gly Asp Lys Leu Arg Met Ile Val
195 200 205

259

Met Gly His Arg Arg Val His Ile Ser Arg Gln Leu Glu Val Glu Pro
 210 215 220

Glu Glu Pro Glu Ala Glu Asn Lys His Lys Pro Arg Arg Lys Ser Lys
 225 230 235 240

Arg Gly Lys Lys Glu Ala Glu Asp Glu Leu Ser Ala Arg His Pro Ala
 245 250 255

Glu Leu Ala Met Glu Pro Thr Pro Glu Leu Pro Ala Glu Val Leu Met
 260 265 270

Val Glu Val Glu Asn Val Val His Glu Asp Phe Gln Val Thr Glu Glu
 275 280 285

Val Lys Ala Leu Thr Ala Glu Ile Val Lys Thr Ile Arg Asp Ile Ile
 290 295 300

Ala Leu Asn Pro Leu Tyr Arg Glu Ser Val Leu Gln Met Met Gln Ala
 305 310 315 320

Gly Gln Arg Val Val Asp Asn Pro Ile Tyr Leu Ser Asp Met Gly Ala
 325 330 335

Ala Leu Thr Gly Ala Glu Ser His Glu Leu Gln Asp Val Leu Glu Glu
 340 345 350

Thr Asn Ile Pro Lys Arg Leu Tyr Lys Ala Leu Ser Leu Leu Lys Lys
 355 360 365

Glu Phe Glu Leu Ser Lys Leu Gln Gln Arg Leu Gly Arg Glu Val Glu
 370 375 380

Glu Lys Ile Lys Gln Thr His Arg Lys Tyr Leu Leu Gln Glu Gln Leu
 385 390 395 400

Lys Ile Ile Lys Lys Glu Leu Gly Leu Glu Lys Asp Asp Lys Asp Ala
 405 410 415

Ile Glu Glu Lys Phe Arg Glu Arg Leu Lys Glu Leu Val Val Pro Lys
 420 425 430

His Val Met Asp Val Val Asp Glu Glu Leu Ser Lys Leu Gly Leu Leu
 435 440 445

260

Asp Asn His Ser Ser Glu Phe Asn Val Thr Arg Asn Tyr Leu Asp Trp
 450 455 460

Leu Thr Ser Ile Pro Trp Gly Lys Tyr Ser Asn Glu Asn Leu Asp Leu
 465 470 475 480

Ala Arg Ala Gln Ala Val Leu Glu Glu Asp His Tyr Gly Met Glu Asp
 485 490 495

Val Lys Lys Arg Ile Leu Glu Phe Ile Ala Val Ser Gln Leu Arg Gly
 500 505 510

Ser Thr Gln Gly Lys Ile Leu Cys Phe Tyr Gly Pro Pro Gly Val Gly
 515 520 525

Lys Thr Ser Ile Ala Arg Ser Ile Ala Arg Ala Leu Asn Arg Glu Tyr
 530 535 540

Phe Arg Phe Ser Val Gly Gly Met Thr Asp Val Ala Glu Ile Lys Gly
 545 550 555 560

His Arg Arg Thr Tyr Val Gly Ala Met Pro Gly Lys Ile Ile Gln Cys
 565 570 575

Leu Lys Lys Thr Lys Thr Glu Asn Pro Leu Ile Leu Ile Asp Glu Val
 580 585 590

Asp Lys Ile Gly Arg Gly Tyr Gln Gly Asp Pro Ser Ser Ala Leu Leu
 595 600 605

Glu Leu Leu Asp Pro Glu Gln Asn Ala Asn Phe Leu Asp His Tyr Leu
 610 615 620

Asp Val Pro Val Asp Leu Ser Lys Val Leu Phe Ile Cys Thr Ala Asn
 625 630 635 640

Val Thr Asp Thr Ile Pro Glu Pro Leu Arg Asp Arg Met Glu Met Ile
 645 650 655

Asn Val Ser Gly Tyr Val Ala Gln Glu Lys Leu Ala Ile Ala Glu Arg
 660 665 670

Tyr Leu Val Pro Gln Ala Arg Ala Leu Cys Gly Leu Asp Glu Ser Lys
 675 680 685

261

Ala Lys Leu Ser Ser Asp Val Leu Thr Leu Leu Ile Lys Gln Tyr Cys
 690 695 700

Arg Glu Ser Gly Val Arg Asn Leu Gln Lys Gln Val Glu Lys Val Leu
 705 710 715 720

Arg Lys Ser Ala Tyr Lys Ile Val Ser Gly Glu Ala Glu Ser Val Glu
 725 730 735

Val Thr Pro Glu Asn Leu Gln Asp Phe Val Gly Lys Pro Val Phe Thr
 740 745 750

Val Glu Arg Met Tyr Asp Val Thr Pro Pro Gly Val Val Met Gly Leu
 755 760 765

Ala Trp Thr Ala Met Gly Glu Arg Gly Gly Gly Arg Arg Pro Gln Ser
 770 775 780

His Ser His Phe Tyr Pro Arg Thr Ser Arg Ser His Leu Val His Leu
 785 790 795 800

Cys Ser Gly Pro Gln Val Ala Leu Asn Gly Leu Trp Arg Gly Val Val
 805 810 815

Cys Trp Glu Pro Arg Gly Ser Gly Leu Arg Glu Gly Lys Ala Val Thr
 820 825 830

His Gly Gly Gly Ser Thr Pro Trp Ala Leu Trp Ile Trp Val Pro Leu
 835 840 845

<210> 189
 <211> 124
 <212> PRT
 <213> Homo sapien

<400> 189

Met Val Phe Leu His Val Gly Gln Ala Gly Leu Glu Leu Pro Thr Ser
 1 5 10 15

Gly Asp Pro Pro Thr Ser Ala Ser Gln Ser Ala Gly Met Thr Glu Leu
 20 25 30

Glu Leu Gly Pro Ser Pro Arg Leu Gln Pro Ile Arg Arg His Pro Lys
 35 40 45

Glu Leu Pro Gln Tyr Gly Gly Pro Gly Lys Asp Ile Phe Glu Asp Gln
 50 55 60

262

Leu Tyr Leu Pro Val His Ser Asp Gly Ile Ser Val His Gln Met Phe
 65 70 75 80

Thr Met Ala Thr Ala Glu His Arg Ser Asn Ser Ser Ile Ala Gly Lys
 85 90 95

Met Leu Thr Lys Val Glu Lys Asn His Glu Lys Glu Lys Ser Gln His
 100 105 110

Leu Glu Gly Ser Ala Ser Ser Ser Leu Ser Ser Asp
 115 120

<210> 190

<211> 296

<212> PRT

<213> Homo sapien

<400> 190

Met Ser Thr Glu Arg Thr Ser Trp Thr Ser Leu Ser Thr Ile Gln Lys
 1 5 10 15

Ile Ala Leu Gly Leu Gly Ile Pro Ala Ser Ala Thr Val Ala Tyr Ile
 20 25 30

Leu Tyr Arg Arg Tyr Arg Glu Ser Arg Glu Glu Arg Leu Thr Phe Val
 35 40 45

Gly Glu Asp Asp Ile Glu Ile Glu Met Arg Val Pro Gln Glu Ala Val
 50 55 60

Lys Leu Ile Ile Gly Arg Gln Gly Ala Asn Ile Lys Gln Leu Arg Lys
 65 70 75 80

Gln Thr Gly Ala Arg Ile Asp Val Asp Thr Glu Asp Val Gly Asp Glu
 85 90 95

Arg Val Leu Leu Ile Ser Gly Phe Pro Val Gln Val Cys Lys Ala Lys
 100 105 110

Ala Ala Ile His Gln Ile Leu Thr Glu Asn Thr Pro Val Ser Glu Gln
 115 120 125

Leu Ser Val Pro Gln Arg Ser Val Gly Arg Ile Ile Gly Arg Gly Gly
 130 135 140

263

Glu Thr Ile Arg Ser Ile Cys Lys Ala Ser Gly Ala Lys Ile Thr Cys
 145 150 155 160

Asp Lys Glu Ser Glu Gly Thr Leu Leu Leu Ser Arg Leu Ile Lys Ile
 165 170 175

Ser Gly Thr Gln Lys Glu Val Ala Ala Ala Lys His Leu Ile Leu Glu
 180 185 190

Lys Val Ser Glu Asp Glu Glu Leu Arg Lys Arg Ile Ala His Ser Ala
 195 200 205

Glu Thr Arg Val Pro Arg Lys Gln Pro Ile Ser Val Arg Arg Glu Asp
 210 215 220

Met Thr Glu Pro Gly Gly Ala Gly Glu Pro Ala Leu Trp Lys Asn Thr
 225 230 235 240

Ser Ser Ser Met Glu Pro Thr Ala Pro Leu Val Thr Pro Pro Pro Lys
 245 250 255

Gly Gly Gly Asp Met Ala Val Val Val Ser Lys Glu Gly Ser Trp Glu
 260 265 270

Lys Pro Ser Asp Asp Ser Phe Gln Lys Ser Glu Ala Gln Ala Ile Pro
 275 280 285

Glu Met Pro Met Phe Glu Ser Met
 290 295

<210> 191

<211> 195

<212> PRT

<213> Homo sapien

<400> 191

Ala Arg Tyr Glu Ala Trp Gly Glu Ser Ala Glu Ala His Val Leu Glu
 1 5 10 15

Gly Pro Asp Thr Asn Thr Thr Ile Ile Gln Leu Gln Pro Leu Gln Glu
 20 25 30

Pro Glu Ser Trp Ala Arg Thr Gln Ser Gly Leu Gln Ser Tyr Leu Leu
 35 40 45

Gln Phe His Gly Leu Val Arg Leu Val His Gln Glu Arg Thr Leu Ala
 50 55 60

264

Phe Pro Leu Thr Ile Arg Cys Phe Leu Gly Cys Glu Leu Pro Pro Glu
65 70 75 80

Gly Ser Arg Ala His Val Phe Phe Glu Val Ala Val Asn Gly Ser Ser
85 90 95

Phe Val Ser Phe Arg Pro Glu Arg Ala Leu Trp Gln Ala Asp Thr Gln
100 105 110

Val Thr Ser Gly Val Val Thr Phe Thr Leu Gln Gln Leu Asn Ala Tyr
115 120 125

Asn Arg Thr Arg Tyr Glu Leu Arg Glu Phe Leu Glu Asp Thr Cys Val
130 135 140

Gln Tyr Val Gln Lys His Ile Ser Ala Glu Asn Thr Lys Gly Ser Gln
145 150 155 160

Thr Ser Arg Ser Tyr Thr Ser Leu Val Leu Gly Val Leu Val Gly Ser
165 170 175

Phe Ile Ile Ala Gly Val Ala Val Gly Ile Phe Leu Cys Thr Gly Gly
180 185 190

Arg Arg Cys
195

<210> 192
<211> 194
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (6)..(7)
<223> X=any amino acid

<400> 192

Leu Gly Thr Gly Arg Xaa Xaa Ser Ala Glu Ala His Val Leu Glu Gly
1 5 10 15

Pro Asp Thr Asn Thr Thr Ile Ile Gln Leu Gln Pro Leu Gln Glu Pro
20 25 30

Glu Ser Trp Ala Arg Thr Gln Ser Gly Leu Gln Ser Tyr Leu Leu Gln
35 40 45

265

Phe His Gly Leu Val Arg Leu Val His Gln Glu Arg Thr Leu Ala Phe
 50 55 60

Pro Leu Thr Ile Arg Cys Phe Leu Gly Cys Glu Leu Pro Pro Glu Gly
 65 70 75 80

Ser Arg Ala His Val Phe Phe Glu Val Ala Val Asn Gly Ser Ser Phe
 85 90 95

Val Ser Phe Arg Pro Glu Arg Ala Leu Trp Gln Ala Asp Thr Gln Val
 100 105 110

Thr Ser Gly Val Val Thr Phe Thr Leu Gln Gln Leu Asn Ala Tyr Asn
 115 120 125

Arg Thr Arg Tyr Glu Leu Arg Glu Phe Leu Glu Asp Thr Cys Val Gln
 130 135 140

Tyr Val Gln Lys His Ile Ser Ala Glu Asn Thr Lys Gly Ser Gln Thr
 145 150 155 160

Ser Arg Ser Tyr Thr Ser Leu Val Leu Gly Val Leu Val Gly Ser Phe
 165 170 175

Ile Ile Ala Gly Val Ala Val Gly Ile Phe Leu Cys Thr Gly Gly Arg
 180 185 190

Arg Cys

<210> 193
 <211> 132
 <212> PRT
 <213> Homo sapien

<400> 193

Met Asn Tyr Thr Gln Arg Glu Leu Gln Met Ala Ala Pro Thr His Leu
 1 5 10 15

Asn His Thr Val Val Gly Thr Pro Cys Gly Asn Gln Thr Leu Ala Thr
 20 25 30

Thr Arg Arg Lys His Leu Ala Trp Arg Glu Arg Arg Pro Ala His Thr
 35 40 45

266

Thr Pro Ala Arg Ala Arg Asp Gly Asn Pro Asn Ile Gly Val Gly Ala
50 55 60

Ala Asp Lys Pro Pro Ser Leu Leu Asn His Ala Arg Arg Ser Ser Leu
65 70 75 80

Pro Asn Arg Pro Pro Arg Ser Thr Gly Gly Asp Glu Ser Leu Ile Thr
85 90 95

His Asn Pro Ser Tyr Ser His Gly Arg Arg Ala Ile Leu His Ala Cys
100 105 110

Val Val Pro His His Thr Glu Arg Arg Val Ala Ser Ile Ile Cys Arg
115 120 125

Pro Gly Pro Arg
130

<210> 194
<211> 199
<212> PRT
<213> Homo sapien

<400> 194

Ile Ile Thr Ile Lys Leu Phe Lys Lys Lys Lys Lys Lys Gln Thr Lys
1 5 10 15

Asn Ile Lys Thr Lys Lys Gln Lys Glu Lys Lys Lys Lys Gln Asn Arg
20 25 30

Gly Ala Gly Ala Pro Gln Lys Lys Ser Arg Pro Arg Gly Ala Ala Pro
35 40 45

Ile Thr Ala Arg Pro Pro Gly Phe Leu Val Ser Thr Gln Gln Gly Gly
50 55 60

Ala Pro His Lys Asn Glu Arg Ala His Arg Ser His Asn Tyr Thr Gln
65 70 75 80

Arg Glu Leu Arg Trp Pro Arg Pro His Thr Ser Thr Thr Gln Ser Trp
85 90 95

Val His Pro Ala Ala Thr Lys Arg Trp Pro Gln Pro Gly Gly Asn Ile
100 105 110

Leu Arg Gly Glu Lys Gly Asp Pro Leu Thr Leu Pro Gln Arg Glu Arg
115 120 125

267

Gly Thr Ala Thr Gln Thr Ser Val Trp Gly Pro Pro Thr Asn Leu Pro
 130 135 140

His Ser Ser Thr Met Arg Asp Asp Leu His Tyr Gln Ile Ala Pro Leu
 145 150 155 160

Ala Arg Arg Ala Ala Thr Asn Arg Ser Ser His Thr Ile His Leu Ile
 165 170 175

His Thr Ala Val Ala Leu Ser Ser Thr Arg Val Trp Cys Leu Ile Thr
 180 185 190

Pro Asn Asp Gly Ser His Gln
 195

<210> 195
 <211> 259
 <212> PRT
 <213> Homo sapien

<400> 195

Met Glu Gly Arg Arg Gln Ala Arg Arg Leu Arg Gln Leu Ala Gly Ala
 1 5 10 15

Gly Ala Gln Gly Gly Ser Pro Ile Val Glu Ala Ala Glu Asn Cys His
 20 25 30

Gly Ala Ala Ser Val Gln Arg Ala Ser Leu Ile Tyr Ser Gly Arg Val
 35 40 45

Cys Ile Trp Gly Gly Val Gln Gly Ala Asp Lys Asp Arg Arg Ala Pro
 50 55 60

Gly Ala Ala Ala Gly Gly Ala Lys Thr Gly Thr Arg Met Ser Ala Pro
 65 70 75 80

Gln Arg Pro Trp Ala Leu Ala Ala Gly Ala Arg Arg Thr Pro Arg Glu
 85 90 95

Ala Gly Ile Ser Lys Ala Val Arg Pro Glu Ala Arg Pro Arg Leu Arg
 100 105 110

Thr Lys Thr Glu Leu Glu Asp Leu Gln Lys Lys Pro Pro Pro Tyr Leu
 115 120 125

268

Arg Asn Leu Ser Ser Asp Asp Ala Asn Val Leu Val Trp His Ala Leu
 130 135 140

Leu Leu Pro Asp Gln Pro Pro Tyr His Leu Lys Ala Phe Asn Leu Arg
 145 150 155 160

Ile Ser Phe Pro Pro Glu Tyr Pro Phe Lys Pro Pro Met Ile Lys Phe
 165 170 175

Thr Thr Lys Ile Tyr His Pro Asn Val Asp Glu Asn Gly Gln Ile Cys
 180 185 190

Leu Pro Ile Ile Ser Ser Glu Asn Trp Lys Pro Cys Thr Lys Thr Cys
 195 200 205

Gln Val Leu Glu Ala Leu Asn Val Leu Val Asn Arg Pro Asn Ile Arg
 210 215 220

Glu Pro Leu Arg Met Asp Leu Ala Asp Leu Leu Thr Gln Asn Pro Glu
 225 230 235 240

Leu Phe Arg Lys Asn Ala Glu Glu Phe Thr Leu Arg Phe Gly Val Asp
 245 250 255

Arg Pro Ser

<210> 196
 <211> 156
 <212> PRT
 <213> Homo sapien

<400> 196

Met Arg Arg Trp Thr Trp Lys Ala Ser Trp Ser Ser Glu Leu Glu Asp
 1 5 10 15

Leu Gln Lys Lys Pro Pro Pro Tyr Leu Arg Asn Leu Ser Ser Asp Asp
 20 25 30

Ala Asn Val Leu Val Trp His Ala Leu Leu Leu Pro Asp Gln Pro Pro
 35 40 45

Tyr His Leu Lys Ala Phe Asn Leu Arg Ile Ser Phe Pro Pro Glu Tyr
 50 55 60

Pro Phe Lys Pro Pro Met Ile Lys Phe Thr Thr Lys Ile Tyr His Pro
 65 70 75 80

269

Asn Val Asp Glu Asn Gly Gln Ile Cys Leu Pro Ile Ile Ser Ser Glu
85 90 95

Asn Trp Lys Pro Cys Thr Lys Thr Cys Gln Val Leu Glu Ala Leu Asn
100 105 110

Val Leu Val Asn Arg Pro Asn Ile Arg Glu Pro Leu Arg Met Asp Leu
115 120 125

Ala Asp Leu Leu Thr Gln Asn Pro Glu Leu Phe Arg Lys Asn Ala Glu
130 135 140

Glu Phe Thr Leu Arg Phe Gly Val Asp Arg Pro Ser
145 150 155

<210> 197
<211> 168
<212> PRT
<213> Homo sapien

<400> 197

Asn Leu Gly Gln Ser Ser Leu Ser Ile Leu Trp Arg Lys Arg Cys Gly
1 5 10 15

Gly Gly Pro Gly Arg Pro Pro Gly Glu Leu Glu Asp Leu Gln Lys Lys
20 25 30

Pro Pro Pro Tyr Leu Arg Asn Leu Ser Ser Asp Asp Ala Asn Val Leu
35 40 45

Val Trp His Ala Leu Leu Leu Pro Asp Gln Pro Pro Tyr His Leu Lys
50 55 60

Ala Phe Asn Leu Arg Ile Ser Phe Pro Pro Glu Tyr Pro Phe Lys Pro
65 70 75 80

Pro Met Ile Lys Phe Thr Thr Lys Ile Tyr His Pro Asn Val Asp Glu
85 90 95

Asn Gly Gln Ile Cys Leu Pro Ile Ile Ser Ser Glu Asn Trp Lys Pro
100 105 110

Cys Thr Lys Thr Cys Gln Val Leu Glu Ala Leu Asn Val Leu Val Asn
115 120 125

270

Arg Pro Asn Ile Arg Glu Pro Leu Arg Met Asp Leu Ala Asp Leu Leu
 130 135 140

Thr Gln Asn Pro Glu Leu Phe Arg Lys Asn Ala Glu Glu Phe Thr Leu
 145 150 155 160

Arg Phe Gly Val Asp Arg Pro Ser
 165

<210> 198
 <211> 137
 <212> PRT
 <213> Homo sapien

<400> 198

Met Glu Gly Arg Arg Gln Gly Asn Leu Ser Ser Asp Asp Ala Asn Val
 1 5 10 15

Leu Val Trp His Ala Leu Leu Leu Pro Asp Gln Pro Pro Tyr His Leu
 20 25 30

Lys Ala Phe Asn Leu Arg Ile Ser Phe Pro Pro Glu Tyr Pro Phe Lys
 35 40 45

Pro Pro Met Ile Lys Phe Thr Thr Lys Ile Tyr His Pro Asn Val Asp
 50 55 60

Glu Asn Gly Gln Ile Cys Leu Pro Ile Ile Ser Ser Glu Asn Trp Lys
 65 70 75 80

Pro Cys Thr Lys Thr Cys Gln Val Leu Glu Ala Leu Asn Val Leu Val
 85 90 95

Asn Arg Pro Asn Ile Arg Glu Pro Leu Arg Met Asp Leu Ala Asp Leu
 100 105 110

Leu Thr Gln Asn Pro Glu Leu Phe Arg Lys Asn Ala Glu Glu Phe Thr
 115 120 125

Leu Arg Phe Gly Val Asp Arg Pro Ser
 130 135

<210> 199
 <211> 237
 <212> PRT
 <213> Homo sapien

<400> 199

271

Met Met Ala Ser Met Arg Val Val Lys Glu Leu Glu Asp Leu Gln Lys
 1 5 10 15
 Lys Pro Pro Pro Tyr Leu Arg Asn Leu Ser Ser Asp Asp Ala Asn Val
 20 25 30
 Leu Val Trp His Ala Leu Leu Leu Pro Glu Ala Glu Val Ala Val Ser
 35 40 45
 Arg Asp His Ala Ile Ala Leu Gln Pro Gly Gln Gln Ser Glu Thr Pro
 50 55 60
 Ser Gln Lys Lys Lys Lys Lys Glu Ala Trp His Gln His Leu Leu Leu
 65 70 75 80
 Met Arg Pro Ser Gly Ser Phe His Ser Trp Trp Lys Ala Lys Gly Ser
 85 90 95
 His Val Tyr Arg Ser His Ala Arg Glu Glu Val Lys Glu Arg Glu Ser
 100 105 110
 Glu Gln Val Pro Gly Ser Ser Lys Gln Pro Ala Phe Ser Asp Gln Pro
 115 120 125
 Pro Tyr His Leu Lys Ala Phe Asn Leu Arg Ile Ser Phe Pro Pro Glu
 130 135 140
 Tyr Pro Phe Lys Pro Pro Met Ile Lys Phe Thr Thr Lys Ile Tyr His
 145 150 155 160
 Pro Asn Val Asp Glu Asn Gly Gln Ile Cys Leu Pro Ile Ile Ser Ser
 165 170 175
 Glu Asn Trp Lys Pro Cys Thr Lys Thr Cys Gln Val Leu Glu Ala Leu
 180 185 190
 Asn Val Leu Val Asn Arg Pro Asn Ile Arg Glu Pro Leu Arg Met Asp
 195 200 205
 Leu Ala Asp Leu Leu Thr Gln Asn Pro Glu Leu Phe Arg Lys Asn Ala
 210 215 220
 Glu Glu Phe Thr Leu Arg Phe Gly Val Asp Arg Pro Ser
 225 230 235

272

<210> 200
 <211> 156
 <212> PRT
 <213> Homo sapien

<400> 200

Gly Pro Gln Glu Ala Ser Thr His Gly Gly Arg Gln Arg Gly Ala Thr
 1 5 10 15

Cys Thr Asp His Ile Ala Arg Glu Glu Val Lys Glu Arg Glu Ser Glu
 20 25 30

Gln Val Pro Gly Ser Ser Lys Gln Pro Ala Phe Ser Asp Gln Pro Pro
 35 40 45

Tyr His Leu Lys Ala Phe Asn Leu Arg Ile Ser Phe Pro Pro Glu Tyr
 50 55 60

Pro Phe Lys Pro Pro Met Ile Lys Phe Thr Thr Lys Ile Tyr His Pro
 65 70 75 80

Asn Val Asp Glu Asn Gly Gln Ile Cys Leu Pro Ile Ile Ser Ser Glu
 85 90 95

Asn Trp Lys Pro Cys Thr Lys Thr Cys Gln Val Leu Glu Ala Leu Asn
 100 105 110

Val Leu Val Asn Arg Pro Asn Ile Arg Glu Pro Leu Arg Met Asp Leu
 115 120 125

Ala Asp Leu Leu Thr Gln Asn Pro Glu Leu Phe Arg Lys Asn Ala Glu
 130 135 140

Glu Phe Thr Leu Arg Phe Gly Val Asp Arg Pro Ser
 145 150 155

<210> 201
 <211> 88
 <212> PRT
 <213> Homo sapien

<400> 201

Met Val Gln Ala Gly Pro Ser Ser Cys Ser Ile Ser Gly Asp Pro Gly
 1 5 10 15

Leu Pro Arg Arg Trp Arg Pro Ala Gln Val Val Arg Pro Gly Arg Leu
 20 25 30

273

Arg Ile Arg Gly Trp Ser Arg Arg Ile Pro Lys Ala Glu Val Gly Ser
 35 40 45

Pro Gly Asp Ser Gln Leu Leu Ser Leu Trp Arg Arg Gly Pro Val Thr
 50 55 60

Glu Ala Pro Phe Ser Asn Pro Gly Ala Ala Phe Gln Arg Leu Asn Phe
 65 70 75 80

Ser Asn His Cys Phe Asn Ser Phe
 85

<210> 202
 <211> 40
 <212> PRT
 <213> Homo sapien

<400> 202

Met Glu Lys Gly Val Gly Gly Gln Pro Arg Gly Arg Arg Arg Ile Tyr
 1 5 10 15

Asn Ile Phe Phe Arg His Arg Cys Tyr Arg Lys Met Cys Glu Arg Ser
 20 25 30

Gly Cys Ala Ala Arg Thr Gly Ala
 35 40

<210> 203
 <211> 60
 <212> PRT
 <213> Homo sapien

<400> 203

Gly Pro Glu Lys Trp Arg Arg Gly Trp Gly Asp Ser His Val Ala Ala
 1 5 10 15

Gly Gly Phe Thr Thr Phe Ser Phe Ala Ile Asp Val Ile Ala Lys Cys
 20 25 30

Val Arg Glu Ala Ala Ala Gln Pro Gly Arg Glu Arg Glu Gly Ala Gly
 35 40 45

Gln Arg Phe Pro Pro Thr Gly Asn Leu Met Gly Leu
 50 55 60

<210> 204
 <211> 213

274

<212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (5)..(5)
 <223> X=any amino acid

<400> 204

Met Pro Gln Asn Xaa Gly Val Ile Gly Leu Arg His His Phe Ala Ile
 1 5 10 15

His Tyr Pro Ala Gly Gly Gly Leu Trp Asp Gly Leu His Gly Val Ala
 20 25 30

Ala Val Gln Gly Ile Thr Lys Ile Lys Val Leu Ala Ile Tyr Ser Phe
 35 40 45

Cys Ser Gln Ile Cys Asp Pro Arg Thr Thr Gly Ala Phe Trp Gln Thr
 50 55 60

Trp Lys Asp Phe Glu Val Arg His Gly Asn Glu Asp Thr Ile Lys Glu
 65 70 75 80

Met Leu Arg Ile Arg Arg Ser Val Gln Ala Thr Tyr Asn Thr Gln Val
 85 90 95

Asn Phe Met Ala Ser Gln Met Leu Lys Val Ser Gly Ser Ala Thr Gly
 100 105 110

Thr Val Ser Asp Leu Ala Pro Gly Gln Ser Gly Met Asp Asp Met Lys
 115 120 125

Leu Leu Glu Gln Arg Ala Glu Gln Leu Ala Ala Glu Ala Glu Arg Asp
 130 135 140

Gln Pro Leu Arg Ala Gln Ser Lys Ile Leu Phe Val Arg Ser Asp Ala
 145 150 155 160

Ser Arg Glu Glu Leu Ala Glu Leu Ala Gln Gln Val Asn Pro Glu Glu
 165 170 175

Ile Gln Leu Gly Glu Asp Glu Asp Glu Asp Glu Met Asp Leu Glu Pro
 180 185 190

Asn Glu Val Arg Leu Glu Gln Gln Ser Val Pro Ala Ala Val Phe Gly
 195 200 205

275

Ser Leu Lys Glu Asp
210

<210> 205
<211> 458
<212> PRT
<213> Homo sapien

<400> 205

Met Ile Asp His Tyr Arg Gly Met Gly Pro Leu Met Leu Leu Glu Arg
1 5 10 15

Arg Ser Val Met Asp Arg Gly Arg Gly Arg Tyr Gln Tyr Ser Pro Gln
20 25 30

Asn Gln His Val Glu Gln Gln Pro His Tyr Thr His Lys Pro Thr Leu
35 40 45

Glu Tyr Ser Pro Phe Pro Ile Pro Pro Gln Ser Pro Ala Tyr Glu Pro
50 55 60

Asn Leu Phe Asp Gly Pro Glu Ser Gln Phe Cys Pro Asn Gln Ser Leu
65 70 75 80

Val Ser Leu Leu Gly Asp Gln Arg Glu Ser Glu Asn Ile Ala Asn Pro
85 90 95

Met Gln Thr Ser Ser Ser Val Gln Gln Gln Asn Asp Ala His Leu His
100 105 110

Ser Phe Ser Met Met Pro Ser Ser Ala Cys Glu Ala Met Val Gly His
115 120 125

Glu Met Ala Ser Asp Ser Ser Asn Thr Ser Leu Pro Phe Ser Asn Met
130 135 140

Gly Asn Pro Met Asn Thr Thr Gln Leu Gly Lys Ser Leu Phe Gln Trp
145 150 155 160

Gln Val Glu Gln Glu Glu Ser Lys Leu Ala Asn Ile Ser Gln Asp Gln
165 170 175

Phe Leu Ser Lys Asp Ala Asp Gly Asp Thr Phe Leu His Ile Ala Val
180 185 190

276

Ala Gln Gly Arg Arg Ala Leu Ser Tyr Val Leu Ala Arg Lys Met Asn
 195 200 205

Ala Leu His Met Leu Asp Ile Lys Glu His Asn Gly Gln Ser Ala Phe
 210 215 220

Gln Val Ala Val Ala Ala Asn Gln His Leu Ile Val Gln Asp Leu Val
 225 230 235 240

Asn Ile Gly Ala Gln Val Asn Thr Thr Asp Cys Trp Gly Arg Thr Pro
 245 250 255

Leu His Val Cys Ala Glu Lys Gly His Ser Gln Val Leu Gln Ala Ile
 260 265 270

Gln Lys Gly Ala Val Gly Ser Asn Gln Phe Val Asp Leu Glu Ala Thr
 275 280 285

Asn Tyr Asp Gly Leu Thr Pro Leu His Cys Ala Val Ile Ala His Asn
 290 295 300

Ala Val Val His Glu Leu Gln Arg Asn Gln Gln Pro His Ser Pro Glu
 305 310 315 320

Val Gln Glu Leu Leu Leu Lys Asn Lys Ser Leu Val Asp Thr Ile Lys
 325 330 335

Cys Leu Ile Gln Met Gly Ala Ala Val Glu Ala Lys Asp Arg Lys Ser
 340 345 350

Gly Arg Thr Ala Leu His Leu Ala Ala Glu Glu Ala Asn Leu Glu Leu
 355 360 365

Ile Arg Leu Phe Leu Glu Leu Pro Ser Cys Leu Ser Phe Val Asn Ala
 370 375 380

Lys Ala Tyr Asn Gly Asn Thr Ala Leu His Val Ala Ala Ser Leu Gln
 385 390 395 400

Tyr Arg Leu Thr Gln Leu Asp Ala Val Arg Leu Leu Met Arg Lys Gly
 405 410 415

Ala Asp Pro Ser Thr Arg Asn Leu Glu Asn Glu Gln Pro Val His Leu
 420 425 430

Val Pro Asp Gly Pro Val Gly Glu Gln Ile Arg Arg Ile Leu Lys Gly

277

435

440

445

Lys Ser Ile Gln Gln Arg Ala Pro Pro Tyr
 450 455

<210> 206

<211> 439

<212> PRT

<213> Homo sapien

<400> 206

Trp Ile Val Val Ala Ala Arg Tyr Gln Tyr Ser Pro Gln Asn Gln His
 1 5 10 15

Val Glu Gln Gln Pro His Tyr Thr His Lys Pro Thr Leu Glu Tyr Ser
 20 25 30

Pro Phe Pro Ile Pro Pro Gln Ser Pro Ala Tyr Glu Pro Asn Leu Phe
 35 40 45

Asp Gly Pro Glu Ser Gln Phe Cys Pro Asn Gln Ser Leu Val Ser Leu
 50 55 60

Leu Gly Asp Gln Arg Glu Ser Glu Asn Ile Ala Asn Pro Met Gln Thr
 65 70 75 80

Ser Ser Ser Val Gln Gln Gln Asn Asp Ala His Leu His Ser Phe Ser
 85 90 95

Met Met Pro Ser Ser Ala Cys Glu Ala Met Val Gly His Glu Met Ala
 100 105 110

Ser Asp Ser Ser Asn Thr Ser Leu Pro Phe Ser Asn Met Gly Asn Pro
 115 120 125

Met Asn Thr Thr Gln Leu Gly Lys Ser Leu Phe Gln Trp Gln Val Glu
 130 135 140

Gln Glu Glu Ser Lys Leu Ala Asn Ile Ser Gln Asp Gln Phe Leu Ser
 145 150 155 160

Lys Asp Ala Asp Gly Asp Thr Phe Leu His Ile Ala Val Ala Gln Gly
 165 170 175

Arg Arg Ala Leu Ser Tyr Val Leu Ala Arg Lys Met Asn Ala Leu His
 180 185 190

278

Met Leu Asp Ile Lys Glu His Asn Gly Gln Ser Ala Phe Gln Val Ala
 195 200 205

Val Ala Ala Asn Gln His Leu Ile Val Gln Asp Leu Val Asn Ile Gly
 210 215 220

Ala Gln Val Asn Thr Thr Asp Cys Trp Gly Arg Thr Pro Leu His Val
 225 230 235 240

Cys Ala Glu Lys Gly His Ser Gln Val Leu Gln Ala Ile Gln Lys Gly
 245 250 255

Ala Val Gly Ser Asn Gln Phe Val Asp Leu Glu Ala Thr Asn Tyr Asp
 260 265 270

Gly Leu Thr Pro Leu His Cys Ala Val Ile Ala His Asn Ala Val Val
 275 280 285

His Glu Leu Gln Arg Asn Gln Gln Pro His Ser Pro Glu Val Gln Glu
 290 295 300

Leu Leu Leu Lys Asn Lys Ser Leu Val Asp Thr Ile Lys Cys Leu Ile
 305 310 315 320

Gln Met Gly Ala Ala Val Glu Ala Lys Asp Arg Lys Ser Gly Arg Thr
 325 330 335

Ala Leu His Leu Ala Ala Glu Glu Ala Asn Leu Glu Leu Ile Arg Leu
 340 345 350

Phe Leu Glu Leu Pro Ser Cys Leu Ser Phe Val Asn Ala Lys Ala Tyr
 355 360 365

Asn Gly Asn Thr Ala Leu His Val Ala Ala Ser Leu Gln Tyr Arg Leu
 370 375 380

Thr Gln Leu Asp Ala Val Arg Leu Leu Met Arg Lys Gly Ala Asp Pro
 385 390 395 400

Ser Thr Arg Asn Leu Glu Asn Glu Gln Pro Val His Leu Val Pro Asp
 405 410 415

Gly Pro Val Gly Glu Gln Ile Arg Arg Ile Leu Lys Gly Lys Ser Ile
 420 425 430

279

Gln Gln Arg Ala Pro Pro Tyr
435

<210> 207
<211> 130
<212> PRT
<213> Homo sapien

<400> 207

Met Gln Pro Leu Trp Leu Cys Trp Ala Leu Trp Val Leu Pro Leu Ala
1 5 10 15

Ser Pro Gly Ala Ala Leu Thr Gly Glu Gln Leu Leu Gly Ser Leu Leu
20 25 30

Arg Gln Leu Gln Leu Lys Glu Val Pro Thr Leu Asp Arg Ala Asp Met
35 40 45

Glu Glu Leu Val Ile Pro Thr His Val Arg Ala Gln Tyr Val Ala Leu
50 55 60

Leu Gln Arg Ser His Gly Asp Arg Ser Arg Gly Lys Arg Phe Ser Gln
65 70 75 80

Ser Phe Arg Glu Val Ala Gly Arg Phe Leu Ala Leu Glu Ala Ser Thr
85 90 95

His Leu Leu Val Phe Gly Met Glu Gln Arg Leu Pro Pro Asn Ser Glu
100 105 110

Leu Val Gln Ala Val Leu Arg Leu Phe Gln Glu Cys Thr Leu Thr Cys
115 120 125

Arg Gly
130

<210> 208
<211> 243
<212> PRT
<213> Homo sapien

<400> 208

Asp Pro Pro Ala Ala Phe Ser Arg Asp Ser Pro Thr Leu Pro Leu Ala
1 5 10 15

Pro Pro Gly Gln His His Ala Ala Pro Val Ala Leu Leu Gly Thr Leu
20 25 30

280

Gly Val Ala Pro Gly Gln Pro Arg Gly Arg Pro Asp Arg Gly Ala Ala
 35 40 45

Pro Gly Gln Pro Ala Ala Ala Ala Ala Gln Arg Gly Ala His Pro
 50 55 60

Gly Gln Gly Arg His Gly Gly Ala Gly His Pro His Pro Arg Glu Gly
 65 70 75 80

Pro Val Arg Gly Pro Ala Ala Ala Gln Pro Arg Gly Pro Leu Pro Arg
 85 90 95

Lys Glu Val Gln Pro Glu Leu Pro Arg Gly Gly Arg Gln Val Pro Gly
 100 105 110

Val Gly Gly Gln His Thr Pro Ala Gly Val Arg His Gly Ala Ala Ala
 115 120 125

Ala Ala Gln Gln Arg Ala Gly Ala Gly Arg Ala Ala Ala Leu Pro Gly
 130 135 140

Met Tyr Ile Asp Leu Gln Gly Met Lys Trp Ala Glu Asn Trp Val Leu
 145 150 155 160

Glu Pro Pro Gly Phe Leu Ala Tyr Glu Cys Val Gly Thr Cys Arg Gln
 165 170 175

Pro Pro Glu Ala Leu Ala Phe Lys Trp Pro Phe Leu Gly Pro Arg Gln
 180 185 190

Cys Ile Ala Ser Glu Thr Asp Ser Leu Pro Met Ile Val Ser Ile Lys
 195 200 205

Glu Gly Gly Arg Thr Arg Pro Gln Val Val Ser Leu Pro Asn Met Arg
 210 215 220

Val Gln Lys Cys Ser Cys Ala Ser Asp Gly Ala Leu Val Pro Arg Arg
 225 230 235 240

Leu Gln Pro

<210> 209
 <211> 204
 <212> PRT
 <213> Homo sapien

281

<400> 209

```

Met Glu Arg Leu Thr Leu Pro Leu Gly Gly Ala Ala Ala Val Asp Glu
1           5           10           15

Tyr Leu Glu Tyr Arg Arg Tyr Lys Gln His Lys Thr Asp Leu Glu Ala
20           25           30

Ile Pro Gln Gln Cys Pro Ile Asp Leu Pro Cys Gln Val Thr Gly Cys
35           40           45

Gln Cys Arg Ala Tyr Leu Tyr Val Pro Leu Asn Gly Ser Gln Pro Ile
50           55           60

Arg Cys Arg Cys Lys His Phe Ala Asp Gln His Ser Ala Ala Pro Gly
65           70           75           80

Phe Thr Cys Asn Thr Cys Ser Lys Cys Ser Gly Phe His Ser Cys Phe
85           90           95

Thr Cys Ala Cys Gly Gln Pro Ala Tyr Ala His Asp Thr Val Val Glu
100          105          110

Thr Lys Gln Glu Arg Leu Ala Gln Glu Lys Pro Val Gly Gln Asp Ile
115          120          125

Pro Tyr Ala Ala Met Gly Gly Leu Thr Gly Phe Ser Ser Leu Ala Glu
130          135          140

Gly Tyr Met Arg Leu Asp Asp Ser Gly Ile Val Gly Thr Ser Ser Gln
145          150          155          160

Val Ser Ser Leu Arg Arg Pro Glu Glu Asp Asp Met Ala Phe Phe Glu
165          170          175

Arg Arg Tyr Gln Glu Arg Met Lys Met Glu Lys Ala Ala Lys Trp Lys
180          185          190

Gly Lys Ala Pro Leu Pro Ser Ala Thr Lys Pro Ser
195          200

```

```

<210> 210
<211> 80
<212> PRT
<213> Homo sapien

```

<400> 210

282

Glu Val Gln Glu Ala Ile Phe Phe Arg Val Cys Gly Ala Arg Ser Val
 1 5 10 15

Val Leu Leu Leu Val Ala Val Arg Leu His Thr Leu Leu Ser Cys Pro
 20 25 30

Leu Glu Gln Pro Ala Gly Thr Glu Trp Ile Leu Glu Glu Gly Val Thr
 35 40 45

Thr Gly Pro Pro Arg Lys Pro Arg Ala Asp Ile Tyr Asn Leu Arg Ser
 50 55 60

Pro Asp Glu Phe Ile Val Gly Gln Asn Gln Ala Leu Ile Glu Pro Gly
 65 70 75 80

<210> 211
 <211> 84
 <212> PRT
 <213> Homo sapien

<400> 211

Glu Gln Gln Pro Ser Pro Ile Asp Ser Thr Glu Thr Thr Arg Asn Gln
 1 5 10 15

Gln Val Arg Pro Thr Thr Ser Arg Asn Lys Arg Arg Ala Ala Ser Gln
 20 25 30

His Ile Ser Lys Ala Thr Arg Pro Thr Ala Lys Arg Gln Ala Ala Asp
 35 40 45

Ser Asp Ile Thr Thr Ser Gly Pro Thr Ala Thr Thr Thr Asp Asp Lys
 50 55 60

Asn Asp Val Cys Glu Asp Thr Pro His Arg Arg Thr Thr Gly Trp His
 65 70 75 80

Gln Arg Asp Leu

<210> 212
 <211> 56
 <212> PRT
 <213> Homo sapien

<400> 212

Pro Leu Trp Arg Arg Leu Leu Leu Gly Ser Arg Leu Leu Leu Pro Cys
 1 5 10 15

283

Asn Arg Asn Trp Arg Trp Asn Met Arg Gly Ala Leu Trp Lys Glu Lys
 20 25 30

Asp Arg Pro Cys Ala Phe Met Lys Val Lys Ile Trp Leu Asn Gln Phe
 35 40 45

His Lys Val Thr Val Tyr Ile Ala
 50 55

<210> 213
 <211> 451
 <212> PRT
 <213> Homo sapien

<400> 213

Met Phe Leu Leu Leu Leu His Leu Gln Ile Lys Trp Arg Ala Thr Ile
 1 5 10 15

Asn Leu Leu Ser Val Thr Glu Asp Gly Leu His Phe Val Glu Tyr Tyr
 20 25 30

Leu Asn Arg Ile Ile His Leu Asp Val Asp Ser Glu Ala Lys Lys Leu
 35 40 45

Leu Gly Leu Gly Gln Lys His Leu Val Met Gly Asp Ile Pro Ala Ala
 50 55 60

Val Asn Ala Phe Gln Glu Ala Ala Ser Leu Leu Gly Lys Lys Tyr Gly
 65 70 75 80

Glu Thr Ala Asn Glu Cys Gly Glu Ala Phe Phe Phe Tyr Gly Lys Ser
 85 90 95

Leu Leu Glu Leu Ala Arg Met Glu Asn Gly Val Leu Gly Asn Ala Leu
 100 105 110

Glu Gly Val His Val Glu Glu Glu Glu Gly Glu Lys Thr Glu Asp Glu
 115 120 125

Ser Leu Val Glu Asn Asn Asp Asn Ile Asp Glu Thr Glu Gly Ser Glu
 130 135 140

Glu Asp Asp Lys Glu Asn Asp Lys Thr Glu Glu Met Pro Asn Asp Ser
 145 150 155 160

284

Val Leu Glu Asn Lys Ser Leu Gln Glu Asn Glu Glu Glu Glu Ile Gly
 165 170 175

Asn Leu Glu Leu Ala Trp Asp Met Leu Asp Leu Ala Lys Ile Ile Phe
 180 185 190

Lys Arg Gln Glu Thr Lys Glu Ala Gln Leu Tyr Ala Ala Gln Ala His
 195 200 205

Leu Lys Leu Gly Glu Val Ser Val Glu Ser Glu Asn Tyr Val Gln Ala
 210 215 220

Val Glu Glu Phe Gln Ser Cys Leu Asn Leu Gln Glu Gln Tyr Leu Glu
 225 230 235 240

Ala His Asp Arg Leu Leu Ala Glu Thr His Tyr Gln Leu Gly Leu Ala
 245 250 255

Tyr Gly Tyr Asn Ser Gln Tyr Asp Glu Ala Val Ala Gln Phe Ser Lys
 260 265 270

Ser Ile Glu Val Ile Glu Asn Arg Met Ala Val Leu Asn Glu Gln Val
 275 280 285

Lys Glu Ala Glu Gly Ser Ser Ala Glu Tyr Lys Lys Glu Ile Glu Glu
 290 295 300

Leu Lys Glu Leu Leu Pro Glu Ile Arg Glu Lys Ile Glu Asp Ala Lys
 305 310 315 320

Glu Ser Gln Arg Ser Gly Asn Val Ala Glu Leu Ala Leu Lys Ala Thr
 325 330 335

Leu Val Glu Ser Ser Thr Ser Gly Phe Thr Pro Gly Gly Gly Gly Ser
 340 345 350

Ser Val Ser Met Ile Ala Ser Arg Lys Pro Thr Asp Gly Ala Ser Ser
 355 360 365

Ser Asn Cys Val Thr Asp Ile Ser His Leu Val Arg Lys Lys Arg Lys
 370 375 380

Pro Glu Glu Glu Ser Pro Arg Lys Asp Asp Ala Lys Lys Ala Lys Gln
 385 390 395 400

Glu Pro Glu Val Asn Gly Gly Ser Gly Asp Ala Val Pro Ser Gly Asn

285

405 410 415

Glu Val Ser Glu Asn Met Glu Glu Glu Ala Glu Asn Gln Ala Glu Ser
420 425 430

Arg Ala Ala Val Glu Gly Thr Val Glu Ala Gly Ala Thr Val Glu Ser
435 440 445

Thr Ala Cys
450

<210> 214
<211> 337
<212> PRT
<213> Homo sapien .

<400> 214

Met Ala His Ala Pro Ala Arg Cys Pro Ser Ala Arg Gly Ser Gly Asp
1 5 10 15

Gly Glu Met Gly Lys Pro Arg Asn Val Ala Leu Ile Thr Gly Ile Thr
20 25 30

Gly Gln Asp Gly Ser Tyr Leu Ala Glu Phe Leu Leu Glu Lys Gly Tyr
35 40 45

Glu Val His Gly Ile Val Arg Arg Ser Ser Ser Phe Asn Thr Gly Arg
50 55 60

Ile Glu His Leu Tyr Lys Asn Pro Gln Ala His Ile Glu Gly Asn Met
65 70 75 80

Lys Leu His Tyr Gly Asp Leu Thr Asp Ser Thr Cys Leu Val Lys Ile
85 90 95

Ile Asn Glu Val Lys Pro Thr Glu Ile Tyr Asn Leu Gly Ala Gln Ser
100 105 110

His Val Lys Ile Ser Phe Asp Leu Ala Glu Tyr Thr Ala Asp Val Asp
115 120 125

Gly Val Gly Thr Leu Arg Leu Leu Asp Ala Val Lys Thr Cys Gly Leu
130 135 140

Ile Asn Ser Val Lys Phe Tyr Gln Ala Ser Thr Ser Glu Leu Tyr Gly
145 150 155 160

286

Lys Val Gln Glu Ile Pro Gln Lys Glu Thr Thr Pro Phe Tyr Pro Arg
 165 170 175

Ser Pro Tyr Gly Ala Asn Phe Val Thr Arg Lys Ile Ser Arg Ser Val
 180 185 190

Ala Lys Ile Tyr Leu Gly Gln Leu Glu Cys Phe Ser Leu Gly Asn Leu
 195 200 205

Asp Ala Lys Arg Asp Trp Gly His Ala Lys Asp Tyr Val Glu Ala Met
 210 215 220

Trp Leu Met Leu Gln Asn Asp Glu Pro Glu Asp Phe Val Ile Ala Thr
 225 230 235 240

Gly Glu Val His Ser Val Arg Glu Phe Val Glu Lys Ser Phe Leu His
 245 250 255

Ile Gly Lys Thr Ile Val Trp Glu Gly Lys Asn Glu Asn Glu Val Gly
 260 265 270

Arg Cys Lys Glu Thr Gly Lys Val His Val Thr Val Asp Leu Lys Tyr
 275 280 285

Tyr Arg Pro Thr Glu Val Asp Phe Leu Gln Gly Asp Cys Thr Lys Ala
 290 295 300

Lys Gln Lys Leu Asn Trp Lys Pro Arg Val Ala Phe Asp Glu Leu Val
 305 310 315 320

Arg Glu Met Val His Ala Asp Val Glu Leu Met Arg Thr Asn Pro Asn
 325 330 335

Ala

<210> 215
 <211> 332
 <212> PRT
 <213> Homo sapien

<400> 215

Met Ala His Ala Pro Ala Arg Cys Pro Ser Ala Arg Gly Ser Gly Asp
 1 5 10 15

Gly Glu Met Gly Lys Pro Arg Asn Val Ala Leu Ile Thr Gly Ile Thr

287

20

25

30

Gly Gln Asp Gly Ser Tyr Leu Ala Glu Phe Leu Leu Glu Lys Gly Tyr
 35 40 45

Glu Val His Gly Ile Val Arg Arg Ser Ser Ser Phe Asn Thr Gly Arg
 50 55 60

Ile Glu His Leu Tyr Lys Asn Pro Gln Ala His Ile Glu Gly Asn Met
 65 70 75 80

Lys Leu His Tyr Gly Asp Leu Thr Asp Ser Thr Cys Leu Val Lys Ile
 85 90 95

Ile Asn Glu Val Lys Pro Thr Glu Ile Tyr Asn Leu Gly Ala Gln Ser
 100 105 110

His Val Lys Ile Ser Phe Asp Leu Ala Glu Tyr Thr Ala Asp Val Asp
 115 120 125

Gly Val Gly Thr Leu Arg Leu Leu Asp Ala Val Lys Thr Cys Gly Leu
 130 135 140

Ile Asn Ser Val Lys Phe Tyr Gln Ala Ser Thr Ser Glu Leu Tyr Gly
 145 150 155 160

Lys Val Gln Glu Ile Pro Gln Lys Glu Thr Thr Pro Phe Tyr Pro Arg
 165 170 175

Ser Pro Tyr Gly Ala Ala Lys Leu Tyr Ala Tyr Trp Ile Val Val Asn
 180 185 190

Phe Arg Glu Ala Tyr Asn Leu Phe Ala Val Asn Gly Ile Leu Phe Asn
 195 200 205

His Glu Ser Pro Arg Arg Gly Ala Asn Phe Val Thr Arg Lys Ile Ser
 210 215 220

Arg Ser Val Ala Lys Ile Tyr Leu Gly Gln Leu Glu Cys Phe Ser Leu
 225 230 235 240

Gly Asn Leu Asp Ala Lys Arg Asp Trp Gly His Ala Lys Asp Tyr Val
 245 250 255

Glu Ala Met Trp Leu Met Leu Gln Asn Asp Glu Pro Glu Asp Phe Val
 260 265 270

288

Ile Ala Thr Gly Glu Val His Ser Val Arg Glu Phe Val Glu Lys Ser
 275 280 285

Phe Leu His Ile Gly Lys Thr Ile Val Trp Glu Gly Lys Asn Glu Asn
 290 295 300

Glu Val Gly Arg Cys Lys Glu Thr Gly Lys Val His Val Thr Val Asp
 305 310 315 320

Leu Lys Tyr Tyr Arg Pro Thr Glu Val Glu Thr Asn
 325 330

<210> 216

<211> 382

<212> PRT

<213> Homo sapien

<400> 216

Met Ala His Ala Pro Ala Arg Cys Pro Ser Ala Arg Gly Ser Gly Asp
 1 5 10 15

Gly Glu Met Gly Lys Pro Arg Asn Val Ala Leu Ile Thr Gly Ile Thr
 20 25 30

Gly Gln Asp Gly Ser Tyr Leu Ala Glu Phe Leu Leu Glu Lys Gly Tyr
 35 40 45

Glu Val His Gly Ile Val Arg Arg Ser Ser Ser Phe Asn Thr Gly Arg
 50 55 60

Ile Glu His Leu Tyr Lys Asn Pro Gln Ala His Ile Glu Gly Asn Met
 65 70 75 80

Lys Leu His Tyr Gly Asp Leu Thr Asp Ser Thr Cys Leu Val Lys Ile
 85 90 95

Ile Asn Glu Val Lys Pro Thr Glu Ile Tyr Asn Leu Gly Ala Gln Ser
 100 105 110

His Val Lys Ile Ser Phe Asp Leu Ala Glu Tyr Thr Ala Asp Val Asp
 115 120 125

Gly Val Gly Thr Leu Arg Leu Leu Asp Ala Val Lys Thr Cys Gly Leu
 130 135 140

289

Ile Asn Ser Val Lys Phe Tyr Gln Ala Ser Thr Ser Glu Leu Tyr Gly
 145 150 155 160

Lys Val Gln Glu Ile Pro Gln Lys Glu Thr Thr Pro Phe Tyr Pro Arg
 165 170 175

Ser Pro Tyr Gly Ala Ala Lys Leu Tyr Ala Tyr Trp Ile Val Val Asn
 180 185 190

Phe Arg Glu Ala Tyr Asn Leu Phe Ala Val Asn Gly Ile Leu Phe Asn
 195 200 205

His Glu Ser Pro Arg Arg Gly Ala Asn Phe Val Thr Arg Lys Ile Ser
 210 215 220

Arg Ser Val Ala Lys Ile Tyr Leu Gly Gln Leu Glu Cys Phe Ser Leu
 225 230 235 240

Gly Asn Leu Asp Ala Lys Arg Asp Trp Gly His Ala Lys Asp Tyr Val
 245 250 255

Glu Ala Met Trp Leu Met Leu Gln Asn Asp Glu Pro Glu Asp Phe Val
 260 265 270

Ile Ala Thr Gly Glu Val His Ser Val Arg Glu Phe Val Glu Lys Ser
 275 280 285

Phe Leu His Ile Gly Lys Thr Ile Val Trp Glu Gly Lys Asn Glu Asn
 290 295 300

Glu Val Gly Arg Cys Lys Glu Thr Gly Lys Val His Val Thr Val Asp
 305 310 315 320

Leu Lys Tyr Tyr Arg Pro Thr Glu Val Val Arg Thr Leu Trp Pro Pro
 325 330 335

Ser Ala Trp Pro Arg Leu Ala Gly Trp Leu Gly Lys Cys Ala His Gly
 340 345 350

Met Pro Gly Ala Ser Leu Trp Ser Cys Gln Phe Ser Ser Leu Ala Ser
 355 360 365

Phe Ser Val His Phe Gln Asn Gln Asn Thr Val Asn Ser Ile
 370 375 380

<210> 217

290

<211> 258

<212> PRT

<213> Homo sapien

<400> 217

Met Asn Ser Asn Val Glu Asn Leu Pro Pro His Ile Ile Arg Leu Val
 1 5 10 15

Tyr Lys Glu Val Thr Thr Leu Thr Ala Asp Pro Pro Asp Gly Ile Lys
 20 25 30

Val Phe Pro Asn Glu Glu Asp Leu Thr Asp Leu Gln Val Thr Ile Glu
 35 40 45

Gly Pro Glu Gly Thr Pro Tyr Ala Gly Gly Leu Phe Arg Met Lys Leu
 50 55 60

Leu Leu Gly Lys Asp Phe Pro Ala Ser Pro Pro Lys Gly Tyr Phe Leu
 65 70 75 80

Thr Lys Ile Phe His Pro Asn Val Gly Ala Asn Gly Glu Ile Cys Val
 85 90 95

Asn Val Leu Lys Arg Asp Trp Thr Ala Glu Leu Gly Ile Arg His Val
 100 105 110

Leu Leu Thr Ile Lys Cys Leu Leu Ile His Pro Asn Pro Glu Ser Ala
 115 120 125

Leu Asn Glu Glu Ala Gly Arg Leu Leu Leu Glu Asn Tyr Glu Glu Tyr
 130 135 140

Ala Ala Arg Ala Arg Leu Leu Thr Glu Ile His Gly Gly Ala Gly Gly
 145 150 155 160

Pro Ser Gly Arg Ala Glu Ala Gly Arg Ala Leu Ala Ser Gly Thr Glu
 165 170 175

Ala Ser Ser Thr Asp Pro Gly Ala Pro Gly Arg Arg Ala Glu Val His
 180 185 190

Trp Pro Glu His Val Gly Arg Glu Arg Arg Trp Gly Arg Lys Thr Thr
 195 200 205

Asp Gly Ala Arg Val Lys Val Phe Leu Ser Arg Asp His Ser Ala Pro
 210 215 220

291

Asn Phe Ser Asn Cys Gly Pro Ser Gly Arg Arg Val Asn Ala Gln Thr
 225 230 235 240

Lys Lys Pro Ser Arg Lys Gly Val Leu Ser Ala Ala Phe Gln Ala Ser
 245 250 255

Leu Leu

<210> 218
 <211> 262
 <212> PRT
 <213> Homo sapien

<400> 218

Met Arg Ala Val Ile Lys Arg Gln Trp Cys Ala Arg Gly Arg Leu Ser
 1 5 10 15

Ala Ala Gly His Arg Gly Gly Gly Leu Val Tyr Ala Val Arg Gly Gly
 20 25 30

Arg Arg Arg Gln Arg Gly Ala Glu Arg Gly Arg Arg Gly Leu Ser Arg
 35 40 45

Ala Ala Ala Ala Ala Val Gly Pro Pro Ala Pro Ala Gly Gly Pro Lys
 50 55 60

Asn Leu Pro Pro His Ile Ile Arg Leu Val Tyr Lys Glu Val Thr Thr
 65 70 75 80

Leu Thr Ala Asp Pro Pro Asp Gly Ile Lys Val Phe Pro Asn Glu Glu
 85 90 95

Asp Leu Thr Asp Leu Gln Val Thr Ile Glu Gly Pro Glu Gly Thr Pro
 100 105 110

Tyr Ala Gly Gly Leu Phe Arg Met Lys Leu Leu Leu Gly Lys Asp Phe
 115 120 125

Pro Ala Ser Pro Pro Lys Gly Tyr Phe Leu Thr Lys Ile Phe His Pro
 130 135 140

Asn Val Gly Ala Asn Gly Glu Ile Cys Val Asn Val Leu Lys Arg Asp
 145 150 155 160

Trp Thr Ala Glu Leu Gly Ile Arg His Val Leu Leu Thr Ile Lys Cys

292

165 170 175

Leu Leu Ile His Pro Asn Pro Glu Ser Ala Leu Asn Glu Glu Ala Gly
 180 185 190

Arg Leu Leu Leu Glu Asn Tyr Glu Glu Tyr Ala Ala Arg Ala Arg Leu
 195 200 205

Leu Thr Glu Ile His Gly Gly Ala Gly Gly Pro Ser Gly Arg Ala Glu
 210 215 220

Ala Gly Arg Ala Leu Ala Ser Gly Thr Glu Ala Ser Ser Thr Asp Pro
 225 230 235 240

Gly Ala Pro Gly Gly Pro Gly Gly Ala Glu Gly Ser His Gly Gln Glu
 245 250 255

Ala Cys Trp Arg Ala Arg
 260

<210> 219
 <211> 291
 <212> PRT
 <213> Homo sapien

<400> 219

Gly Ser Glu Leu Arg Gly Arg Gly Gly Arg Gly Leu Arg Ala Val Ile
 1 5 10 15

Lys Arg Gln Trp Cys Ala Arg Gly Arg Leu Ser Ala Ala Gly His Arg
 20 25 30

Gly Gly Gly Leu Val Tyr Ala Val Arg Gly Gly Arg Arg Arg Gln Arg
 35 40 45

Gly Ala Glu Arg Gly Arg Arg Gly Leu Ser Arg Ala Ala Ala Ala Ala
 50 55 60

Val Gly Pro Pro Ala Pro Ala Gly Gly Pro Lys Asn Leu Pro Pro His
 65 70 75 80

Ile Ile Arg Leu Val Tyr Lys Glu Val Thr Thr Leu Thr Ala Asp Pro
 85 90 95

Pro Asp Gly Ile Lys Val Phe Pro Asn Glu Glu Asp Leu Thr Asp Leu
 100 105 110

293

Gln Val Thr Ile Glu Gly Pro Glu Gly Thr Pro Tyr Ala Gly Gly Leu
 115 120 125

Phe Arg Met Lys Leu Leu Leu Gly Lys Asp Phe Pro Ala Ser Pro Pro
 130 135 140

Lys Gly Tyr Phe Leu Thr Lys Ile Phe His Pro Asn Val Gly Ala Asn
 145 150 155 160

Gly Glu Ile Cys Val Asn Val Leu Lys Arg Asp Trp Thr Ala Glu Leu
 165 170 175

Gly Ile Arg His Val Leu Leu Thr Ile Lys Cys Leu Leu Ile His Pro
 180 185 190

Asn Pro Glu Ser Ala Leu Asn Glu Glu Ala Gly Arg Leu Leu Leu Glu
 195 200 205

Asn Tyr Glu Glu Tyr Ala Ala Arg Ala Arg Leu Leu Thr Glu Ile His
 210 215 220

Gly Gly Ala Gly Gly Pro Ser Gly Arg Ala Glu Ala Gly Arg Ala Leu
 225 230 235 240

Ala Ser Gly Thr Glu Ala Ser Ser Thr Asp Pro Gly Ala Pro Gly Gly
 245 250 255

Pro Gly Gly Ala Glu Gly Pro Met Ala Lys Lys His Ala Gly Glu Arg
 260 265 270

Asp Lys Lys Leu Ala Ala Lys Lys Lys Thr Asp Lys Lys Arg Ala Leu
 275 280 285

Arg Arg Leu
 290

<210> 220
 <211> 233
 <212> PRT
 <213> Homo sapien

<400> 220

Met Asn Ser Asn Val Glu Asn Leu Pro Pro His Ile Ile Arg Leu Val
 1 5 10 15

Tyr Lys Glu Val Thr Thr Leu Thr Ala Asp Pro Pro Asp Gly Ile Lys

294

20

25

30

Val Phe Pro Asn Glu Glu Asp Leu Thr Asp Leu Gln Val Thr Ile Glu
 35 40 45

Gly Pro Glu Gly Thr Pro Tyr Ala Gly Gly Leu Phe Arg Met Lys Leu
 50 55 60

Leu Leu Gly Lys Asp Phe Pro Ala Ser Pro Pro Lys Gly Tyr Phe Leu
 65 70 75 80

Thr Lys Ile Phe His Pro Asn Val Gly Ala Asn Gly Glu Ile Cys Val
 85 90 95

Asn Val Leu Lys Arg Asp Trp Thr Ala Glu Leu Gly Ile Arg His Val
 100 105 110

Leu Leu Leu Ser Trp Lys Asp Lys Gln Cys Gln Thr Gln Asp Thr Gln
 115 120 125

Val Leu Leu Arg Ser Ala Gln Glu His Leu Thr Met Gln Arg Val Thr
 130 135 140

Ile Lys Cys Leu Leu Ile His Pro Asn Pro Glu Ser Ala Leu Asn Glu
 145 150 155 160

Glu Ala Gly Arg Leu Leu Leu Glu Asn Tyr Glu Glu Tyr Ala Ala Arg
 165 170 175

Ala Arg Leu Leu Thr Glu Ile His Gly Gly Ala Gly Gly Pro Ser Gly
 180 185 190

Arg Ala Glu Ala Gly Arg Ala Leu Ala Ser Gly Thr Glu Ala Ser Ser
 195 200 205

Thr Asp Pro Gly Ala Pro Gly Gly Pro Gly Gly Ala Glu Gly Ser His
 210 215 220

Gly Gln Glu Ala Cys Trp Arg Ala Arg
 225 230

<210> 221
 <211> 390
 <212> PRT
 <213> Homo sapien
 <400> 221

295

Glu Pro Ser Arg Pro Pro Arg Ala Pro Ile Gly Arg Pro Ala Thr Gln
 1 5 10 15

Pro Ser Pro Pro Leu Pro Ser Leu Leu Thr Ile Val Cys Gly Glu Gly
 20 25 30

Ser Glu Arg Val Glu Asn Gln Gly Thr Cys Ala Leu Thr Ser Arg Leu
 35 40 45

Arg Leu Gly Ala Glu Gly Pro Arg Arg Ala Trp Pro Ala Gly Gly Tyr
 50 55 60

Lys Glu Ala Val Val Arg Ala Arg Pro Ala Gln Cys Cys Arg Ala Pro
 65 70 75 80

Gly Arg Arg Val Gly Leu Arg Cys Ala Arg Arg Thr Ser Glu Ala Ala
 85 90 95

Gly Ser Gly Ala Gly Pro Pro Gly Pro Leu Gln Gly Arg Ser Gly Ser
 100 105 110

Ser Trp Ala Pro Arg Pro Gly Arg Arg Thr Glu Glu Arg Arg Lys Gly
 115 120 125

Ala Gly Gly Thr Arg Pro Arg Pro Ala Ala Ala Met Asn Ser Asn Val
 130 135 140

Glu Asn Leu Pro Pro His Ile Ile Arg Leu Val Tyr Lys Glu Val Thr
 145 150 155 160

Thr Leu Thr Ala Asp Pro Pro Asp Gly Ile Lys Val Phe Pro Asn Glu
 165 170 175

Glu Asp Leu Thr Asp Leu Gln Val Thr Ile Glu Gly Pro Glu Gly Thr
 180 185 190

Pro Tyr Ala Gly Gly Leu Phe Arg Met Lys Leu Leu Leu Gly Lys Asp
 195 200 205

Phe Pro Ala Ser Pro Pro Lys Gly Tyr Phe Leu Thr Lys Ile Phe His
 210 215 220

Pro Asn Val Gly Ala Asn Gly Glu Ile Cys Val Asn Val Leu Lys Arg
 225 230 235 240

Trp Arg Cys Gln Arg Ser Glu Gly Leu Gln Glu Gly Leu Lys Leu Pro
50 55 60

297

Ala Gln Asn Leu Arg Met Glu Pro Ala Leu His Tyr Leu Arg Ser Gln
65 70 75 80

Gly Leu Gly Arg Trp Arg Lys Val Ile Ser Pro Ser Leu Lys Ser Tyr
85 90 95

Phe Leu Asn Val Ala Pro His Gln Ala Leu Tyr Leu Thr Ser
100 105 110

<210> 223
<211> 257
<212> PRT
<213> Homo sapien

<400> 223

Met Asp His Arg Ser Arg Leu Arg Gly Thr Gly Leu Asn Arg Ile Pro
1 5 10 15

Gly Thr Gln Ser Arg Ala Pro Arg Val Pro Leu Pro Phe His Val Gln
20 25 30

Gln Glu Ala Arg Glu Gly Glu Asp Trp Glu Arg Glu Pro Pro Arg Gln
35 40 45

Arg Pro Pro Ile Tyr Glu Pro Pro Glu Ser Glu Glu Leu Pro Asp Asn
50 55 60

Val Met Val Ser Lys Pro Ala Pro Tyr Trp Glu Gly Thr Ala Val Ile
65 70 75 80

Asp Gly Glu Phe Lys Glu Leu Lys Leu Thr Asp Tyr Arg Gly Lys Tyr
85 90 95

Leu Val Phe Phe Phe Tyr Pro Leu Asp Phe Thr Phe Val Cys Pro Thr
100 105 110

Glu Ile Ile Ala Phe Gly Asp Arg Leu Glu Glu Phe Arg Ser Ile Asn
115 120 125

Thr Glu Val Val Ala Cys Ser Val Asp Ser Gln Phe Thr His Leu Ala
130 135 140

Trp Ile Asn Thr Pro Arg Arg Gln Gly Gly Leu Gly Pro Ile Arg Ile
145 150 155 160

298

Pro Leu Leu Ser Asp Leu Thr His Gln Ile Ser Lys Asp Tyr Gly Val
 165 170 175

Tyr Leu Glu Asp Ser Gly His Thr Leu Arg Gly Leu Phe Ile Ile Asp
 180 185 190

Asp Lys Gly Ile Leu Arg Gln Ile Thr Leu Asn Asp Leu Pro Val Gly
 195 200 205

Arg Ser Val Asp Glu Thr Leu Arg Leu Val Gln Ala Phe Gln Tyr Thr
 210 215 220

Asp Lys His Gly Glu Val Cys Pro Ala Gly Trp Lys Pro Gly Ser Glu
 225 230 235 240

Thr Ile Ile Pro Asp Pro Ala Gly Lys Leu Lys Tyr Phe Asp Lys Leu
 245 250 255

Asn

<210> 224
 <211> 105
 <212> PRT
 <213> Homo sapien

<400> 224

Met Gln Lys Lys Lys Asn Ser Asn Ser Asn Ser Gly Thr Ser Ser Phe
 1 5 10 15

Gly Lys Arg Arg Asn Lys Thr His Thr Leu Cys Arg Arg Cys Gly Ser
 20 25 30

Lys Ala Tyr His Leu Gln Lys Ser Thr Cys Gly Lys Cys Gly Tyr Pro
 35 40 45

Ala Lys Arg Lys Arg Lys Tyr Asn Trp Ser Ala Lys Ala Lys Arg Arg
 50 55 60

Asn Thr Thr Gly Thr Gly Arg Met Arg His Leu Lys Ile Val Tyr Arg
 65 70 75 80

Arg Phe Arg His Gly Phe Arg Glu Gly Thr Thr Pro Lys Pro Lys Arg
 85 90 95

Ala Ala Val Ala Ala Ser Ser Ser Ser
 100 105

299

<210> 225
 <211> 111
 <212> PRT
 <213> Homo sapien

<400> 225

Ile Phe Met Val Gly Val Asp Ala Lys Lys Lys Glu Phe Glu Phe Glu
 1 5 10 15

Phe Gly Thr Ser Ser Phe Gly Lys Arg Arg Asn Lys Thr His Thr Leu
 20 25 30

Cys Arg Arg Cys Gly Ser Lys Ala Tyr His Leu Gln Lys Ser Thr Cys
 35 40 45

Gly Lys Cys Gly Tyr Pro Ala Lys Arg Lys Arg Lys Tyr Asn Trp Ser
 50 55 60

Ala Lys Ala Lys Arg Arg Asn Thr Thr Gly Thr Gly Arg Met Arg His
 65 70 75 80

Leu Lys Ile Val Tyr Arg Arg Phe Arg His Gly Phe Arg Glu Gly Thr
 85 90 95

Thr Pro Lys Pro Lys Arg Ala Ala Val Ala Ala Ser Ser Ser Ser
 100 105 110

<210> 226
 <211> 104
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (3)..(5)
 <223> X=any amino acid

<400> 226

Met Ser Xaa Xaa Xaa Arg Ile Arg Pro Arg Gly Thr Ser Ser Phe Gly
 1 5 10 15

Lys Arg Arg Asn Lys Thr His Thr Leu Cys Arg Arg Cys Gly Ser Lys
 20 25 30

Ala Tyr His Leu Gln Lys Ser Thr Cys Gly Lys Cys Gly Tyr Pro Ala
 35 40 45

300

Lys Arg Lys Arg Lys Tyr Asn Trp Ser Ala Lys Ala Lys Arg Arg Asn
 50 55 60

Thr Thr Gly Thr Gly Arg Met Arg His Leu Lys Ile Val Tyr Arg Arg
 65 70 75 80

Phe Arg His Gly Phe Arg Glu Gly Thr Thr Pro Lys Pro Lys Arg Ala
 85 90 95

Ala Val Ala Ala Ser Ser Ser Ser
 100

<210> 227
 <211> 129
 <212> PRT
 <213> Homo sapien

<220>
 <221> MISC_FEATURE
 <222> (12)..(12)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (25)..(25)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (62)..(62)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (64)..(64)
 <223> X=any amino acid

<400> 227

Gln Ser His Lys Thr Leu Val Leu Gln Thr Thr Xaa Arg Ser Leu Leu
 1 5 10 15

Ala His Thr Thr Cys Arg Phe Trp Xaa Phe Pro Asn Leu Leu Gly Ile
 20 25 30

Lys Val Asn Asn Ser Ile Thr Arg Gly Ser Gly Gln Pro Ser Phe Val
 35 40 45

Arg Gly Cys Ile Val Gly Lys Pro Thr Ser Val Cys Gln Xaa Leu Xaa

301

50

55

60

Glu Phe Gly Arg Gly Glu Arg His Arg Leu Glu Ser Val Ala Ile Arg
65 70 75 80

Arg Thr Arg Cys Ala Ala Ala Val Ala Leu Arg Pro Thr Thr Phe Arg
85 90 95

Ser Arg Pro Val Ala Asn Val Ala Thr Leu Pro Ser Ala Arg Glu Ser
100 105 110

Ile Thr Gly Val Pro Arg Leu Lys Asp Glu Ile Pro Pro Glu Leu Val
115 120 125

Glu

<210> 228

<211> 96

<212> PRT

<213> Homo sapien

<400> 228

Ala Cys Arg Ala Ala Gln Cys Asp Gly Ser Trp Ser Arg Pro Arg Ser
1 5 10 15

Leu Cys Arg Arg Cys Gly Ser Lys Ala Tyr His Leu Gln Lys Ser Thr
20 25 30

Cys Gly Lys Cys Gly Tyr Pro Ala Lys Arg Lys Arg Lys Tyr Asn Trp
35 40 45

Ser Ala Lys Ala Lys Arg Arg Asn Thr Thr Gly Thr Gly Arg Met Arg
50 55 60

His Leu Lys Ile Val Tyr Arg Arg Phe Arg His Gly Phe Arg Glu Gly
65 70 75 80

Thr Thr Pro Lys Pro Lys Arg Ala Ala Val Ala Ala Ser Ser Ser Ser
85 90 95

<210> 229

<211> 55

<212> PRT

<213> Homo sapien

<400> 229

302

Met His Ala Glu Arg Arg Ser Val Met Asp Arg Gly Arg Gly Arg Gly
 1 5 10 15

Arg Pro Thr Thr Phe Arg Ser Arg Pro Val Ala Asn Val Ala Thr Leu
 20 25 30

Pro Ser Ala Arg Glu Ser Ile Thr Gly Val Pro Arg Leu Lys Asp Glu
 35 40 45

Ile Pro Pro Glu Leu Val Glu
 50 55

<210> 230
 <211> 72
 <212> PRT
 <213> Homo sapien

<400> 230

Ala Tyr His Leu Gln Lys Ser Thr Cys Gly Lys Cys Gly Tyr Pro Ala
 1 5 10 15

Lys Arg Lys Arg Lys Tyr Asn Trp Ser Ala Lys Ala Lys Arg Arg Asn
 20 25 30

Thr Thr Gly Thr Gly Arg Met Arg His Leu Lys Ile Val Tyr Arg Arg
 35 40 45

Phe Arg His Gly Phe Arg Glu Gly Thr Thr Pro Lys Pro Lys Arg Ala
 50 55 60

Ala Val Ala Ala Ser Ser Ser
 65 70

<210> 231
 <211> 185
 <212> PRT
 <213> Homo sapien

<400> 231

Met Leu Glu Arg Arg Ser Val Asp Gly Cys Ala Arg Ala Gly Gly Arg
 1 5 10 15

Ala Gly Gly Ala Ile Met Gly Val Asp Ile Arg His Asn Lys Asp Arg
 20 25 30

Lys Val Arg Arg Lys Glu Pro Lys Ser Gln Asp Ile Tyr Leu Arg Leu
 35 40 45

303

Leu Val Lys Leu Tyr Arg Phe Leu Ala Arg Arg Thr Asn Ser Thr Phe
 50 55 60

Asn Gln Val Val Leu Lys Arg Leu Phe Met Ser Arg Thr Asn Arg Pro
 65 70 75 80

Pro Leu Ser Leu Ser Arg Met Ile Arg Lys Met Lys Leu Pro Gly Arg
 85 90 95

Glu Asn Lys Thr Ala Val Val Val Gly Thr Ile Thr Asp Asp Val Arg
 100 105 110

Val Gln Glu Val Pro Lys Leu Lys Val Cys Ala Leu Arg Val Thr Ser
 115 120 125

Arg Ala Arg Ser Arg Ile Leu Arg Ala Gly Gly Lys Ile Leu Thr Phe
 130 135 140

Asp Gln Leu Ala Leu Asp Ser Pro Lys Gly Cys Gly Thr Val Leu Leu
 145 150 155 160

Ser Gly Pro Arg Lys Gly Arg Glu Val Tyr Arg His Phe Gly Lys Ala
 165 170 175

Pro Gly Thr Pro His Ser His Thr Lys
 180 185

<210> 232
 <211> 214
 <212> PRT
 <213> Homo sapien

<400> 232

Gly Leu Trp His Cys Pro Ala Leu Arg Ser Ser Gln Gly Pro Arg Gly
 1 5 10 15

Val Pro Ala Phe Arg Gln Gly Pro Arg Asn Pro Ala Gln Pro His Gln
 20 25 30

Val Ser Ile Arg Pro Pro Ala Leu Pro Ser Pro Gln Thr Gln Pro Ala
 35 40 45

Gly Pro Gly Leu Ala Thr Leu Gly Leu Leu Leu Leu Ser Leu Val Pro
 50 55 60

Ala Ser Pro Arg Pro Ser Gly Thr Leu Ser Cys Leu Ile Leu Pro Ala

[illegible]

```
<210> 233
<211> 131
<212> PRT
<213> Homo sapien
```

<400> 233

Met Leu Glu Arg Arg Ser Val Asp Gly Cys Ala Arg Ala Gly Gly Arg
1 5 10 15

Ala Gly Gly Ala Ile Met Gly Val Asp Ile Arg His Asn Lys Asp Arg
20 25 30

Lys Val Arg Arg Lys Glu Pro Lys Ser Gln Asp Ile Tyr Leu Arg Leu
35 40 45

Leu Val Lys Leu Tyr Arg Phe Leu Ala Arg Arg Thr Asn Ser Thr Phe
50 55 60

305

Asn Gln Val Val Leu Lys Arg Leu Phe Met Ser Arg Thr Asn Arg Pro
 65 70 75 80

Pro Leu Ser Leu Ser Arg Met Ile Arg Lys Met Lys Leu Pro Gly Arg
 85 90 95

Glu Asn Lys Thr Ala Val Val Val Gly Thr Ile Thr Asp Asp Val Arg
 100 105 110

Val Gln Glu Val Pro Lys Leu Ile Gly Arg Asp His Ala Lys Pro Asp
 115 120 125

Ser Ser Thr
 130

<210> 234
 <211> 132
 <212> PRT
 <213> Homo sapien

<400> 234

Asp Ala Cys Ser Ser Gly Ala Gly Asp Gly Cys Ala Arg Ala Gly Gly
 1 5 10 15

Arg Ala Gly Gly Ala Ile Met Gly Val Asp Ile Arg His Asn Lys Asp
 20 25 30

Arg Lys Val Arg Arg Lys Glu Pro Lys Ser Gln Asp Ile Tyr Leu Arg
 35 40 45

Leu Leu Val Lys Leu Tyr Arg Phe Leu Ala Arg Arg Thr Asn Ser Thr
 50 55 60

Phe Asn Gln Val Val Leu Lys Arg Leu Phe Met Ser Arg Thr Asn Arg
 65 70 75 80

Pro Pro Leu Ser Leu Ser Arg Met Ile Arg Lys Met Lys Leu Pro Gly
 85 90 95

Arg Glu Asn Lys Thr Ala Val Val Val Gly Thr Ile Thr Asp Asp Val
 100 105 110

Arg Val Gln Glu Val Pro Lys Leu Ile Gly Arg Asp His Ala Lys Pro
 115 120 125

Asp Ser Ser Thr

306

130

<210> 235
 <211> 195
 <212> PRT
 <213> Homo sapien

<400> 235

Met Asp Trp Ser Arg Arg Gly Gly Arg Ala Gly Gly Ala Ile Met Gly
 1 5 10 15

Val Asp Ile Arg His Asn Lys Asp Arg Lys Val Arg Arg Lys Glu Pro
 20 25 30

Lys Ser Gln Asp Ile Tyr Leu Arg Leu Leu Val Lys Leu Tyr Arg Phe
 35 40 45

Leu Ala Arg Arg Thr Asn Ser Thr Phe Asn Gln Val Val Leu Lys Arg
 50 55 60

Leu Phe Met Ser Arg Thr Asn Arg Pro Pro Leu Ser Leu Ser Arg Met
 65 70 75 80

Val Ser Gly Trp Ser Arg Glu His Gly Arg Pro Gly Ser Arg Trp Val
 85 90 95

Leu Ser Val Trp Lys Gly Gly Arg Thr Trp Ser Ser Gly Ser Asn Gln
 100 105 110

Gly Ile Lys Gly Leu Ser Gln Pro Val Ala Ser Val Glu Leu Gly Leu
 115 120 125

Leu Val Gly Thr Glu Cys Pro Trp Ala Val Gly Lys Ser Pro Gly Pro
 130 135 140

Pro Leu Leu Leu Phe Val Arg Trp Arg Cys Pro Gly Gly Phe Arg Arg
 145 150 155 160

Leu Pro Gln Val Ile Thr Glu Phe Tyr Val Lys Gly Ser Ala Glu Gly
 165 170 175

Gly Pro Ile Glu Gln Ser Ala Phe Phe Phe Leu Ser Gly Ala Phe Pro
 180 185 190

Ser Trp Thr
 195

307

<210> 236
 <211> 115
 <212> PRT
 <213> Homo sapien

<400> 236

Ser Ala Trp Thr Leu Thr Ser Asn Gly Arg Ser Phe Pro Gly Pro Phe
 1 5 10 15

Pro Lys Ser Trp Ala Cys Phe Leu Leu Pro Leu Ala Ile Leu Cys Pro
 20 25 30

Cys Gly Cys Ser Pro Thr Leu Arg Ala Val Pro Asp Leu Ser Cys Tyr
 35 40 45

Phe Pro Lys Pro Glu Thr Ala Pro Leu Gln Ile Leu Ala Ala Pro Phe
 50 55 60

Pro Cys Val Gln Tyr Arg Val Ile Thr Cys Pro Ser Leu Val Pro Leu
 65 70 75 80

Ile Leu Pro Cys Asp Tyr Ser Val Ile Pro Leu Pro Val Pro Glu Pro
 85 90 95

Pro Gly Leu Phe Leu Gly Ser Pro Glu Cys Ser Pro Arg Thr Gln Ser
 100 105 110

Ala Val Pro
 115

<210> 237
 <211> 156
 <212> PRT
 <213> Homo sapien

<400> 237

Gln Ser Leu Gly Arg Gly Leu Ala Thr Thr Arg Gly Arg Gly Ser Asp
 1 5 10 15

Gly Asn Gly Pro Thr Gly Asn Gly Asp His Pro Asn Phe Ser Leu Ser
 20 25 30

Glu Gly Arg Ala Phe Gly Ser Leu Ala Ala Gln Pro Ile Thr Ser Cys
 35 40 45

Leu Ser Val Pro Ala Pro Pro Phe Ser Leu Gly Lys Leu Gln Asp Gly
 50 55 60

308

Leu Leu His Ile Thr Thr Cys Ser Phe Val Ala Pro Trp Asn Ser Leu
65 70 75 80

Ser Leu Ala Gln Arg Arg Gly Phe Thr Lys Thr Tyr Thr Val Gly Cys
85 90 95

Glu Glu Cys Thr Val Phe Pro Cys Leu Ser Ile Pro Cys Lys Leu Gln
100 105 110

Ser Gly Thr His Cys Leu Trp Thr Asp Gln Leu Leu Gln Gly Ser Glu
115 120 125

Lys Gly Phe Gln Ser Arg His Leu Ala Cys Leu Pro Arg Glu Pro Gly
130 135 140

Leu Cys Thr Trp Gln Ser Leu Arg Ser Gln Ile Ala
145 150 155

<210> 238
<211> 226
<212> PRT
<213> Homo sapien

<400> 238

Met Ala Ala Ala Ala Ala Ala Ala Gly Ala Ala Gly Ser Ala Ala Pro
1 5 10 15

Ala Ala Ala Ala Gly Ala Pro Gly Ser Gly Gly Ala Pro Ser Gly Ser
20 25 30

Gln Gly Val Leu Ile Gly Asp Arg Leu Tyr Ser Gly Val Leu Ile Thr
35 40 45

Leu Glu Asn Cys Leu Leu Pro Asp Asp Lys Leu Arg Phe Thr Pro Ser
50 55 60

Met Ser Ser Gly Leu Asp Thr Asp Thr Glu Thr Asp Leu Arg Val Val
65 70 75 80

Gly Cys Glu Leu Ile Gln Ala Ala Gly Ile Leu Leu Arg Leu Pro Gln
85 90 95

Val Ala Met Ala Thr Gly Gln Val Leu Phe Gln Arg Phe Phe Tyr Thr
100 105 110

309

Lys Ser Phe Val Lys His Ser Met Glu His Val Ser Met Ala Cys Val
 115 120 125

His Leu Ala Ser Lys Ile Glu Glu Ala Pro Arg Arg Ile Arg Asp Val
 130 135 140

Ile Asn Val Phe His Arg Leu Arg Gln Leu Arg Asp Lys Lys Lys Pro
 145 150 155 160

Val Pro Leu Leu Leu Asp Gln Asp Tyr Val Asn Leu Lys Asn Gln Ile
 165 170 175

Ile Lys Ala Glu Arg Arg Val Leu Lys Glu Leu Gly Phe Cys Val His
 180 185 190

Val Lys His Pro His Lys Ile Ile Val Met Tyr Leu Gln Val Leu Glu
 195 200 205

Cys Glu Arg Asn Gln His Leu Val Gln Thr Ser Trp Val Ala Ser Glu
 210 215 220

Gly Lys
 225

<210> 239
 <211> 253
 <212> PRT
 <213> Homo sapien

<400> 239

Asp Ser Gln Asp Cys Leu Ala Leu Ser Pro Ser Asn Arg Leu Leu Arg
 1 5 10 15

Gly Val Val Arg Leu Ser Arg Phe Ser Leu Asp Asn Ala Gly Gly Arg
 20 25 30

Pro Gly Phe Pro Gly Gly Ala Leu Gln Phe Phe Leu Cys Leu Ala Ser
 35 40 45

Arg Asn Tyr Met Asn Asp Ser Leu Arg Thr Asp Val Phe Val Arg Phe
 50 55 60

Gln Pro Glu Ser Ile Ala Cys Ala Cys Ile Tyr Leu Ala Ala Arg Thr
 65 70 75 80

Leu Glu Ile Pro Leu Pro Asn Arg Pro His Trp Phe Leu Leu Phe Gly
 85 90 95

310

Ala Thr Glu Glu Glu Ile Gln Glu Ile Cys Leu Lys Ile Leu Gln Leu
 100 105 110

Tyr Ala Arg Lys Lys Val Asp Leu Thr His Leu Glu Gly Glu Val Glu
 115 120 125

Lys Arg Lys His Ala Ile Glu Glu Ala Lys Ala Gln Ala Arg Gly Leu
 130 135 140

Leu Pro Gly Gly Thr Gln Val Leu Asp Gly Thr Ser Gly Phe Ser Pro
 145 150 155 160

Ala Pro Lys Leu Val Glu Ser Pro Lys Glu Gly Lys Gly Ser Lys Pro
 165 170 175

Ser Pro Leu Ser Val Lys Asn Thr Lys Arg Arg Leu Glu Gly Ala Lys
 180 185 190

Lys Ala Lys Ala Asp Ser Pro Val Asn Gly Leu Pro Lys Gly Arg Glu
 195 200 205

Ser Arg Ser Arg Ser Arg Ser Arg Glu Gln Ser Tyr Ser Arg Ser Pro
 210 215 220

Ser Arg Ser Ala Ser Pro Lys Arg Arg Lys Ser Asp Ser Gly Ser Thr
 225 230 235 240

Ser Gly Gly Ser Lys Ser Gln Arg Ser Leu Gln Arg Leu
 245 250

<210> 240

<211> 346

<212> PRT

<213> Homo sapien

<400> 240

Asp Ser Gln Asp Cys Leu Ala Leu Ser Pro Ser Asn Arg Leu Leu Arg
 1 5 10 15

Gly Val Val Arg Leu Ser Arg Phe Ser Leu Asp Asn Ala Gly Gly Arg
 20 25 30

Pro Gly Phe Pro Gly Gly Ala Leu Gln Phe Phe Leu Cys Leu Ala Ser
 35 40 45

311

Arg Asn Tyr Met Asn Asp Ser Leu Arg Thr Asp Val Phe Val Arg Phe
 50 55 60

Gln Pro Glu Ser Ile Ala Cys Ala Cys Ile Tyr Leu Ala Ala Arg Thr
 65 70 75 80

Leu Glu Ile Pro Leu Pro Asn Arg Pro His Trp Phe Leu Leu Phe Gly
 85 90 95

Ala Thr Glu Glu Glu Ile Gln Glu Ile Cys Leu Lys Ile Leu Gln Leu
 100 105 110

Tyr Ala Arg Lys Lys Val Asp Leu Thr His Leu Glu Gly Glu Val Glu
 115 120 125

Lys Arg Lys His Ala Ile Glu Glu Ala Lys Ala Gln Ala Arg Gly Leu
 130 135 140

Leu Pro Gly Gly Thr Gln Val Leu Asp Gly Thr Ser Gly Phe Ser Pro
 145 150 155 160

Ala Pro Lys Leu Val Glu Ser Pro Lys Glu Gly Lys Gly Ser Lys Pro
 165 170 175

Ser Pro Leu Ser Val Lys Asn Thr Lys Arg Arg Leu Glu Gly Ala Lys
 180 185 190

Lys Ala Lys Ala Asp Ser Pro Val Asn Gly Leu Pro Lys Gly Arg Glu
 195 200 205

Ser Arg Ser Arg Ser Arg Ser Arg Glu Gln Ser Tyr Ser Arg Ser Pro
 210 215 220

Ser Arg Ser Ala Ser Pro Lys Arg Arg Lys Ser Asp Ser Gly Ser Thr
 225 230 235 240

Ser Gly Gly Ser Lys Ser Gln Ser Arg Ser Arg Ser Arg Ser Asp Ser
 245 250 255

Pro Pro Arg Gln Ala Pro Arg Ser Ala Pro Tyr Lys Gly Ser Glu Ile
 260 265 270

Arg Gly Ser Arg Lys Ser Lys Asp Cys Lys Tyr Pro Gln Lys Pro His
 275 280 285

Lys Ser Arg Ser Arg Ser Ser Ser Arg Ser Arg Ser Arg Ser Arg Glu

312

290

295

300

Arg Ala Asp Asn Pro Gly Lys Tyr Lys Lys Lys Ser His Tyr Tyr Arg
 305 310 315 320

Asp Gln Arg Arg Glu Arg Ser Arg Ser Tyr Glu Arg Thr Gly Arg Arg
 325 330 335

Tyr Glu Arg Asp His Pro Val Ala Ala Leu
 340 345

<210> 241
 <211> 91
 <212> PRT
 <213> Homo sapien

<400> 241

Pro Thr Thr Thr Lys Phe Ala Ala Ala Ser Thr Phe Leu Asn Trp Cys
 1 5 10 15

Cys Leu Gly Phe Ile Ala Phe Ala Tyr Ser Val Lys Ser Arg Asp Arg
 20 25 30

Lys Met Val Gly Asp Val Thr Gly Ala Gln Ala Tyr Ala Ser Thr Ala
 35 40 45

Lys Cys Leu Asn Ile Trp Ala Leu Ile Leu Gly Ile Phe Met Thr Ile
 50 55 60

Gly Phe Ile Leu Leu Leu Val Phe Gly Ser Val Thr Val Tyr His Ile
 65 70 75 80

Met Leu Gln Ile Ile Gln Glu Lys Arg Gly Tyr
 85 90

<210> 242
 <211> 92
 <212> PRT
 <213> Homo sapien

<400> 242

Gly Gln Glu Asp Gly Trp Arg Arg Asp Arg Gly Pro Gly Leu Cys Leu
 1 5 10 15

His Arg Gln Val Pro Glu His Leu Gly Pro Asp Phe Gly His Leu His
 20 25 30

313

Asp Arg Ile His Pro Val Thr Gly Ile Arg Leu Cys Asp Ser Leu Pro
 35 40 45

Tyr Tyr Val Thr Asp Asn Thr Gly Lys Thr Gly Leu Leu Val Ala Ala
 50 55 60

His Ser Leu Gln Pro Leu His Ser Thr Val Gln Cys Trp Pro Cys Thr
 65 70 75 80

Leu Gly Leu Leu Pro Leu Pro Pro Trp Ser Cys Pro
 85 90

<210> 243

<211> 137

<212> PRT

<213> Homo sapien

<400> 243

Met Leu Glu Arg Arg Ser Val Met Asp Arg Pro Pro Ala Glu Val Arg
 1 5 10 15

Glu Thr Lys Ile Lys Gly Lys Ser Gly Arg Phe Phe Thr Val Lys Leu
 20 25 30

Pro Val Ala Leu Asp Pro Gly Ala Lys Ile Ser Val Ile Val Glu Thr
 35 40 45

Val Tyr Thr His Val Leu His Pro Tyr Pro Thr Gln Ile Thr Gln Ser
 50 55 60

Glu Lys Gln Phe Val Val Phe Glu Gly Asn His Tyr Phe Tyr Ser Pro
 65 70 75 80

Tyr Pro Thr Lys Thr Gln Thr Met Arg Val Lys Leu Ala Ser Arg Asn
 85 90 95

Val Glu Ser Tyr Thr Lys Leu Gly Asn Pro Thr Arg Ser Glu Asp Leu
 100 105 110

Leu Asp Tyr Gly Pro Phe Arg Asp Val Pro Ala Tyr Ser Gln Asp Thr
 115 120 125

Phe Lys Val Pro Arg Pro Arg Pro Arg
 130 135

<210> 244

<211> 148

314

<212> PRT
 <213> Homo sapien
 <220>
 <221> MISC_FEATURE
 <222> (21)..(22)
 <223> X=any amino acid

<220>
 <221> MISC_FEATURE
 <222> (24)..(24)
 <223> X=any amino acid

<400> 244

Arg Leu Ile Tyr Arg Ala Ile Gly His Leu Ile Met Leu Glu Arg Arg
 1 5 10 15

Ser Val Met Asp Xaa Xaa Pro Xaa Glu Val Arg Glu Thr Lys Ile Lys
 20 25 30

Gly Lys Ser Gly Arg Phe Phe Thr Val Lys Leu Pro Val Ala Leu Asp
 35 40 45

Pro Gly Ala Lys Ile Ser Val Ile Val Glu Thr Val Tyr Thr His Val
 50 55 60

Leu His Pro Tyr Pro Thr Gln Ile Thr Gln Ser Glu Lys Gln Phe Val
 65 70 75 80

Val Phe Glu Gly Asn His Tyr Phe Tyr Ser Pro Tyr Pro Thr Lys Thr
 85 90 95

Gln Thr Met Arg Val Lys Leu Ala Ser Arg Asn Val Glu Ser Tyr Thr
 100 105 110

Lys Leu Gly Asn Pro Thr Arg Ser Glu Asp Leu Leu Asp Tyr Gly Pro
 115 120 125

Phe Arg Asp Val Pro Ala Tyr Ser Gln Asp Thr Phe Lys Val Pro Arg
 130 135 140

Pro Arg Pro Arg
 145

<210> 245
 <211> 479
 <212> PRT
 <213> Homo sapien

315

. <400> 245 .

Met Glu Ala Pro Ala Ala Gly Leu Phe Leu Leu Leu Leu Gly Thr
 1 5 10 15

Trp Ala Pro Ala Pro Gly Ser Ala Ser Ser Glu Ala Pro Pro Leu Ile
 20 25 30

Asn Glu Asp Val Lys Arg Thr Val Asp Leu Ser Ser His Leu Ala Lys
 35 40 45

Val Thr Ala Glu Val Val Leu Ala His Leu Gly Gly Gly Ser Thr Ser
 50 55 60

Arg Ala Thr Ser Phe Leu Leu Ala Leu Glu Pro Glu Leu Glu Ala Arg
 65 70 75 80

Leu Ala His Leu Gly Val Gln Val Lys Gly Glu Asp Glu Glu Glu Asn
 85 90 95

Asn Leu Glu Val Arg Glu Thr Lys Ile Lys Gly Lys Ser Gly Arg Phe
 100 105 110

Phe Thr Val Lys Leu Pro Val Ala Leu Asp Pro Gly Ala Lys Ile Ser
 115 120 125

Val Ile Val Glu Thr Val Tyr Thr His Val Leu His Pro Tyr Pro Thr
 130 135 140

Gln Ile Thr Gln Ser Glu Lys Gln Phe Val Val Phe Glu Gly Asn His
 145 150 155 160

Tyr Phe Tyr Ser Pro Tyr Pro Thr Lys Thr Gln Thr Met Arg Val Lys
 165 170 175

Leu Ala Ser Arg Asn Val Glu Ser Tyr Thr Lys Leu Gly Asn Pro Thr
 180 185 190

Arg Ser Glu Asp Leu Leu Asp Tyr Gly Pro Phe Arg Asp Val Pro Ala
 195 200 205

Tyr Ser Gln Asp Thr Phe Lys Val His Tyr Glu Asn Asn Ser Pro Phe
 210 215 220

Leu Thr Ile Thr Ser Met Thr Arg Val Ile Glu Val Ser His Trp Gly
 225 230 235 240

316

Asn Ile Ala Val Glu Glu Asn Val Asp Leu Lys His Thr Gly Ala Val
245 250 255

Leu Lys Gly Pro Phe Ser Arg Tyr Asp Tyr Gln Arg Gln Pro Asp Ser
260 265 270

Gly Ile Ser Ser Ile Arg Ser Phe Lys Thr Ile Leu Pro Ala Ala Ala
275 280 285

Gln Asp Val Tyr Tyr Arg Asp Glu Ile Gly Asn Val Ser Thr Ser His
290 295 300

Leu Leu Ile Leu Asp Asp Ser Val Glu Met Glu Ile Arg Pro Arg Phe
305 310 315 320

Pro Leu Phe Gly Gly Trp Lys Thr His Tyr Ile Val Gly Tyr Asn Leu
325 330 335

Pro Ser Tyr Glu Tyr Leu Tyr Asn Leu Gly Asp Gln Tyr Ala Leu Lys
340 345 350

Met Arg Phe Val Asp His Val Phe Asp Glu Gln Val Ile Asp Ser Leu
355 360 365

Thr Val Lys Ile Ile Leu Pro Glu Gly Ala Lys Asn Ile Glu Ile Asp
370 375 380

Ser Pro Tyr Glu Ile Ser Arg Ala Pro Asp Glu Leu His Tyr Thr Tyr
385 390 395 400

Leu Asp Thr Phe Gly Arg Pro Val Ile Val Ala Tyr Lys Lys Asn Leu
405 410 415

Val Glu Gln His Ile Gln Asp Ile Val Leu Asp Ala Gln Val Lys Glu
420 425 430

Leu Val Leu Lys Ser Ala Val Glu Ala Glu Arg Leu Val Ala Gly Lys
435 440 445

Leu Lys Lys Asp Thr Tyr Ile Glu Asn Glu Lys Leu Ile Ser Gly Lys
450 455 460

Arg Gln Glu Leu Val Thr Lys Ile Asp His Ile Leu Asp Ala Leu
465 470 475

317

<210> 246
 <211> 361
 <212> PRT
 <213> Homo sapien

<400> 246

Met Glu Ala Pro Ala Ala Gly Leu Phe Leu Leu Leu Leu Gly Thr
 1 5 10 15

Trp Ala Pro Ala Pro Gly Ser Ala Ser Ser Glu Ala Pro Pro Leu Ile
 20 25 30

Asn Glu Asp Val Lys Arg Thr Val Asp Leu Ser Ser His Leu Ala Lys
 35 40 45

Val Thr Ala Glu Val Val Leu Ala His Leu Gly Gly Gly Ser Thr Ser
 50 55 60

Arg Ala Thr Ser Phe Leu Leu Ala Leu Glu Pro Glu Leu Glu Ala Arg
 65 70 75 80

Leu Ala His Leu Gly Val Gln Val Lys Gly Glu Asp Glu Glu Glu Asn
 85 90 95

Asn Leu Glu Val Arg Glu Thr Lys Ile Lys Gly Lys Ser Gly Arg Phe
 100 105 110

Phe Thr Val Lys Leu Pro Val Ala Leu Asp Pro Gly Ala Lys Ile Ser
 115 120 125

Val Ile Val Glu Thr Val Tyr Thr His Val Leu His Pro Tyr Pro Thr
 130 135 140

Gln Ile Thr Gln Ser Glu Lys Gln Phe Val Val Phe Glu Gly Asn His
 145 150 155 160

Tyr Phe Tyr Ser Pro Tyr Pro Thr Lys Thr Gln Thr Met Arg Val Lys
 165 170 175

Leu Ala Ser Arg Asn Val Glu Ser Tyr Thr Lys Leu Gly Asn Pro Thr
 180 185 190

Arg Ser Glu Asp Leu Leu Asp Tyr Gly Pro Phe Arg Asp Val Pro Ala
 195 200 205

Tyr Ser Gln Asp Thr Phe Lys Val His Tyr Glu Asn Asn Ser Pro Phe

210

220

Asn Ile Ala Val Glu Glu Asn Val Asp Leu Lys His Thr Gly Ala Val
245 250 255

Leu Lys Gly Pro Phe Ser Arg Tyr Asp Tyr Gln Arg Gln Pro Asp Ser
260 265 270

Gly Ile Ser Ser Ile Arg Ser Phe Lys Thr Ile Leu Pro Ala Ala Ala
275 280 285

Gln Asp Val Tyr Tyr Arg Asp Glu Ile Gly Asn Val Ser Thr Ser His
290 295 300

Leu Leu Ile Leu Asp Asp Ser Val Glu Met Glu Ile Arg Pro Arg Phe
305 310 315 320

Pro Leu Phe Gly Gly Trp Lys Thr His Tyr Ile Val Gly Tyr Asn Leu
325 330 335

Pro Ser Tyr Glu Tyr Leu Tyr Asn Leu Gly Gln Ser Ser Ile Val Arg
340 345 350

Glu Lys Leu Thr Phe Ser Leu Ile Ser
355 360

```
<210> 247
<211> 420
<212> PRT
<213> Homo sapien
```

<400> 247

Met Glu Ala Pro Ala Ala Gly Leu Phe Leu Leu Leu Leu Gly Thr
1 5 10 15

Trp Ala Pro Ala Pro Gly Ser Ala Ser Ser Glu Ala Pro Pro Leu Ile
20 25 30

Asn Glu Asp Val Lys Arg Thr Val Asp Leu Ser Ser His Leu Ala Lys
35 40 45

Val Thr Ala Glu Val Val Leu Ala His Leu Gly Gly Gly Ser Thr Ser
50 55 60

319

Arg Ala Thr Ser Phe Leu Leu Ala Leu Glu Pro Glu Leu Glu Ala Arg
 65 70 75 80

Leu Ala His Leu Gly Val Gln Val Lys Gly Glu Asp Glu Glu Glu Asn
 85 90 95

Asn Leu Glu Val Arg Glu Thr Lys Ile Lys Gly Lys Ser Gly Arg Phe
 100 105 110

Phe Thr Val Lys Leu Pro Val Ala Leu Asp Pro Gly Ala Lys Ile Ser
 115 120 125

Val Ile Val Glu Thr Val Tyr Thr His Val Leu His Pro Tyr Pro Thr
 130 135 140

Gln Ile Thr Gln Ser Glu Lys Gln Phe Val Val Phe Glu Gly Asn His
 145 150 155 160

Tyr Phe Tyr Ser Pro Tyr Pro Thr Lys Thr Gln Thr Met Arg Val Lys
 165 170 175

Leu Ala Ser Arg Asn Val Glu Ser Tyr Thr Lys Leu Gly Asn Pro Thr
 180 185 190

Arg Ser Glu Asp Leu Leu Asp Tyr Gly Pro Phe Arg Asp Val Pro Ala
 195 200 205

Tyr Ser Gln Asp Thr Phe Lys Val His Tyr Glu Asn Asn Ser Pro Phe
 210 215 220

Leu Thr Ile Thr Ser Met Thr Arg Val Ile Glu Val Ser His Trp Gly
 225 230 235 240

Asn Ile Ala Val Glu Glu Asn Val Asp Leu Lys His Thr Gly Ala Val
 245 250 255

Leu Lys Gly Pro Phe Ser Arg Tyr Asp Tyr Gln Arg Gln Pro Asp Ser
 260 265 270

Gly Ile Ser Ser Ile Arg Ser Phe Lys Thr Ile Leu Pro Ala Ala Ala
 275 280 285

Gln Asp Val Tyr Tyr Arg Asp Glu Ile Gly Asn Val Ser Thr Ser His
 290 295 300

320

Leu Leu Ile Leu Asp Asp Ser Val Glu Met Glu Ile Arg Pro Arg Phe
 305 310 315 320

Pro Leu Phe Gly Gly Trp Lys Thr His Tyr Ile Val Gly Tyr Asn Leu
 325 330 335

Pro Ser Tyr Glu Tyr Leu Tyr Asn Leu Gly Asp Gln Tyr Ala Leu Lys
 340 345 350

Met Arg Phe Val Asp His Val Phe Asp Glu Gln Val Ile Asp Ser Leu
 355 360 365

Thr Val Lys Ile Ile Leu Pro Glu Gly Ala Lys Thr Thr Ser Trp Met
 370 375 380

Pro Cys Ser Pro Cys Pro His Pro Pro Gly Gly Pro Gly Cys Leu His
 385 390 395 400

Phe Ala Val Ala Gly Arg Leu Gly Gly Ser Gly Arg Leu Cys Met Glu
 405 410 415

Ala Ser Glu Ser
 420

<210> 248
 <211> 128
 <212> PRT
 <213> Homo sapien

<400> 248

Gly Cys Ala Arg Glu Pro Glu Ser Arg Leu Pro Lys Leu Gly Ser Trp
 1 5 10 15

Glu Asn Leu Gly Pro Gly Leu Thr Glu Lys Arg Arg Gly Lys Glu Ala
 20 25 30

Gly Gln Glu Glu Gly Ala Trp Arg Thr Pro Ala Gly Gly Arg Gly Ala
 35 40 45

Ala Gly Leu Ser Val Thr Pro Leu Ser Pro Pro Arg Pro Ala Pro Pro
 50 55 60

Ala Gly Glu Gly Pro Arg Cys Pro Pro Gly Arg Pro Ala Pro Ala Arg
 65 70 75 80

Arg Arg Lys Gly Trp Arg Val Glu Arg Arg Gly Gly Arg Gly Ser Ala
 85 90 95

321

Trp Asp Ala Pro Gly His Arg Ala Arg Ser Leu Arg Pro Gly Ala Gly
 100 105 110

Gln Val Arg Gly Gln Asp Val Gly Arg Thr Trp Cys Met Ala Ala Ser
 115 120 125

<210> 249

<211> 315

<212> PRT

<213> Homo sapien

<400> 249

Met Ser Ala Ala Gly Ala Gly Ala Gly Val Glu Ala Gly Phe Ser Ser
 1 5 10 15

Glu Glu Leu Leu Ser Leu Arg Phe Pro Leu His Arg Ala Cys Arg Asp
 20 25 30

Gly Asp Leu Ala Thr Leu Cys Ser Leu Leu Gln Gln Thr Pro His Ala
 35 40 45

His Leu Ala Ser Glu Asp Ser Phe Tyr Gly Trp Thr Pro Val His Trp
 50 55 60

Ala Ala His Phe Gly Lys Leu Glu Cys Leu Val Gln Leu Val Arg Ala
 65 70 75 80

Gly Ala Thr Leu Asn Val Ser Thr Thr Arg Tyr Ala Gln Thr Pro Ala
 85 90 95

His Ile Ala Ala Phe Gly Gly His Pro Gln Cys Leu Val Trp Leu Ile
 100 105 110

Gln Ala Gly Ala Asn Ile Asn Lys Pro Asp Cys Glu Gly Glu Thr Pro
 115 120 125

Ile His Lys Ala Ala Arg Ser Gly Ser Leu Glu Cys Ile Ser Ala Leu
 130 135 140

Val Ala Asn Gly Ala His Val Asp Leu Arg Asn Ala Ser Gly Leu Thr
 145 150 155 160

Ala Ala Asp Ile Ala Gln Thr Gln Gly Phe Gln Glu Cys Ala Gln Phe
 165 170 175

322

Leu Leu Asn Leu Gln Asn Cys His Leu Asn His Phe Tyr Asn Asn Gly
 180 185 190

Ile Leu Asn Gly Gly His Gln Asn Val Phe Pro Asn His Ile Ser Val
 195 200 205

Gly Thr Asn Arg Lys Arg Cys Leu Glu Asp Ser Glu Asp Phe Gly Val
 210 215 220

Lys Lys Ala Arg Thr Glu Ala Gln Ser Leu Asp Ser Ala Val Pro Leu
 225 230 235 240

Thr Asn Gly Asp Thr Glu Asp Asp Ala Asp Lys Met His Val Asp Arg
 245 250 255

Glu Phe Ala Val Val Thr Gly Gly Ser Gly Gln Phe Pro Val Ser Cys
 260 265 270

Asn Asn Asn Pro Met Val Glu Asp Thr Lys Gln Gln Glu Ser Gly Ser
 275 280 285

Val Gly Pro Lys Glu Ile Glu Ile Tyr Thr Val Ser Ala Met Gln Thr
 290 295 300

Pro Cys Arg Cys Arg Asn Gln Tyr Glu Lys Gln
 305 310 315

<210> 250
 <211> 142
 <212> PRT
 <213> Homo sapien

<400> 250

Met Gly Asn Phe Lys Ser Ile Ser Thr Ser Thr Lys Met Val Asn Gly
 1 5 10 15

Arg Lys Ile Thr Thr Lys Arg Ile Val Glu Asn Gly Gln Glu Arg Val
 20 25 30

Glu Val Glu Glu Asp Gly Gln Leu Lys Ser Leu Thr Ile Asn Gly Val
 35 40 45

Ala Asp Asp Asp Ala Leu Ala Glu Glu Arg Met Arg Arg Gly Gln Asn
 50 55 60

Ala Leu Pro Ala Gln Pro Ala Gly Leu Arg Pro Pro Lys Pro Pro Arg
 65 70 75 80

323

Pro Ala Ser Leu Leu Arg His Ala Pro His Cys Leu Ser Glu Glu Glu
85 90 95

Gly Glu Gln Asp Arg Pro Arg Ala Pro Gly Pro Trp Asp Pro Leu Ala
100 105 110

Ser Ala Ala Gly Leu Lys Glu Gly Gly Lys Arg Lys Lys Gln Lys Gln
115 120 125

Arg Glu Glu Ser Lys Lys Lys Lys Ser Thr Lys Gly Asn His
130 135 140

<210> 251
<211> 72
<212> PRT
<213> Homo sapien

<400> 251

Met Gly Leu Ser His Ala Gly Trp His Arg Ala Gly Lys His Glu Ala
1 5 10 15

Ser Pro His Gln Gly Phe Ala Cys Arg Lys Ala Ala Leu Trp Pro Ala
20 25 30

Gly Glu Ala Glu Glu Thr Pro Val Asp Thr Leu Pro Thr Gly Leu Lys
35 40 45

Glu Gly Gly Lys Arg Lys Lys Gln Lys Gln Arg Glu Glu Ser Lys Lys
50 55 60

Lys Lys Ser Thr Lys Gly Asn His
65 70

<210> 252
<211> 122
<212> PRT
<213> Homo sapien

<400> 252

Thr Gly Leu Glu Ala Arg Gly Ala Pro Pro Asp Ala Gly Ala Pro Pro
1 5 10 15

Cys Ser Ala Cys Gly Arg Ala His Ala Leu Gly Ser Ser Val Gly Gln
20 25 30

Asp Cys Leu Glu Ala Thr Leu Ala Arg Gln Asp Tyr Ala Ile Thr Asp

324

35

40

45

Gln Ser Glu Gln Gly Gln Glu Thr Gly Leu Thr Ala Arg Val Ala Gly
 50 55 60

Thr Asp Val Trp Asp Leu Ala Ala Thr Leu Cys Phe Ser Pro Ala Leu
 65 70 75 80

Asn Leu Leu His Phe Pro Leu Val Leu Pro Asp Pro Leu His Ser Phe
 85 90 95

Arg Leu Leu Asn His Ser Ala Cys Cys Trp Asn Ile Ser Gly Phe Arg
 100 105 110

Ser Thr Gly Gly Arg Arg Trp Leu Thr Glu
 115 120

<210> 253
 <211> 42
 <212> PRT
 <213> Homo sapien

<400> 253

Met Ala Lys Lys Ala Gly Leu Cys Leu Gly Gly Ser Arg Gln Gly Gly
 1 5 10 15

Cys Gln Ser Gly Met Val Thr Gly Asn Glu Pro Arg Asp Leu Ala Leu
 20 25 30

Ser His Pro Leu Ser Phe Val Gly Gly Leu
 35 40

<210> 254
 <211> 260
 <212> PRT
 <213> Homo sapien

<400> 254

Val Phe Cys Ser Phe Phe Ala Glu Lys Glu Gln Gln Glu Ala Ile Glu
 1 5 10 15

His Ile Asp Glu Val Gln Asn Glu Ile Asp Arg Leu Asn Glu Gln Ala
 20 25 30

Ser Glu Glu Ile Leu Lys Val Glu Gln Lys Tyr Asn Lys Leu Arg Gln
 35 40 45

325

```

Pro Phe Phe Gln Lys Arg Ser Glu Leu Ile Ala Lys Ile Pro Asn Phe
 50          55          60

Trp Val Thr Thr Phe Val Asn His Pro Gln Val Ser Ala Leu Leu Gly
65          70          75          80

Glu Glu Asp Glu Glu Ala Leu His Tyr Leu Thr Arg Val Glu Val Thr
      85          90          95

Glu Phe Glu Asp Ile Lys Ser Gly Tyr Arg Ile Asp Phe Tyr Phe Asp
      100          105          110

Glu Asn Pro Tyr Phe Glu Asn Lys Val Leu Ser Lys Glu Phe His Leu
      115          120          125

Asn Glu Ser Gly Asp Pro Ser Ser Lys Ser Thr Glu Ile Lys Trp Lys
      130          135          140

Ser Gly Lys Asp Leu Thr Lys Arg Ser Ser Gln Thr Gln Asn Lys Ala
145          150          155          160

Ser Arg Lys Arg Gln His Glu Glu Pro Glu Ser Phe Phe Thr Trp Phe
      165          170          175

Thr Asp His Ser Asp Ala Gly Ala Asp Glu Leu Gly Glu Val Ile Lys
      180          185          190

Asp Asp Ile Trp Pro Asn Pro Leu Gln Tyr Tyr Leu Val Pro Asp Met
      195          200          205

Asp Asp Glu Glu Gly Glu Gly Glu Glu Asp Asp Asp Asp Asp Glu Glu
      210          215          220

Glu Glu Gly Leu Glu Asp Ile Asp Glu Glu Gly Asp Glu Asp Glu Gly
225          230          235          240

Glu Glu Asp Glu Asp Asp Asp Glu Gly Glu Glu Gly Glu Glu Asp Glu
      245          250          255

Gly Glu Asp Asp
      260

```

<210> 255

<211> 285

<212> PRT

<213> Homo sapien

326

<400> 255

Ser Leu Gln Asp Lys Arg Ala Pro Ile Pro Glu His Thr Pro Phe Ser
 1 5 10 15

Ser Ser Pro Phe Cys Ala Ser Leu Leu Ser Asp Leu Ile Val Ala Pro
 20 25 30

Lys Lys Glu Gln Gln Glu Ala Ile Glu His Ile Asp Glu Val Gln Asn
 35 40 45

Glu Ile Asp Arg Leu Asn Glu Gln Ala Ser Glu Glu Ile Leu Lys Val
 50 55 60

Glu Gln Lys Tyr Asn Lys Leu Arg Gln Pro Phe Phe Gln Lys Arg Ser
 65 70 75 80

Glu Leu Ile Ala Lys Ile Pro Asn Phe Trp Val Thr Thr Phe Val Asn
 85 90 95

His Pro Gln Val Ser Ala Leu Leu Gly Glu Glu Asp Glu Glu Ala Leu
 100 105 110

His Tyr Leu Thr Arg Val Glu Val Thr Glu Phe Glu Asp Ile Lys Ser
 115 120 125

Gly Tyr Arg Ile Asp Phe Tyr Phe Asp Glu Asn Pro Tyr Phe Glu Asn
 130 135 140

Lys Val Leu Ser Lys Glu Phe His Leu Asn Glu Ser Gly Asp Pro Ser
 145 150 155 160

Ser Lys Ser Thr Glu Ile Lys Trp Lys Ser Gly Lys Asp Leu Thr Lys
 165 170 175

Arg Ser Ser Gln Thr Gln Asn Lys Ala Ser Arg Lys Arg Gln His Glu
 180 185 190

Glu Pro Glu Ser Phe Phe Thr Trp Phe Thr Asp His Ser Asp Ala Gly
 195 200 205

Ala Asp Glu Leu Gly Glu Val Ile Lys Asp Asp Ile Trp Pro Asn Pro
 210 215 220

Leu Gln Tyr Tyr Leu Val Pro Asp Met Asp Asp Glu Glu Gly Glu Gly
 225 230 235 240

Glu Glu Asp Asp Asp Asp Asp Glu Glu Glu Glu Gly Leu Glu Asp Ile
245 250 255

Asp Glu Glu Gly Asp Glu Asp Glu Gly Glu Glu Asp Glu Asp Asp Asp
260 265 270

Glu Gly Glu Glu Gly Glu Glu Asp Glu Gly Glu Asp Asp
275 280 285

<210>	256
<211>	600
<212>	PRT
<213>	Homo sapien

<400> 256

Met Ala Thr Pro Leu Pro Gly Arg Ala Gly Gly Pro Ala Thr Pro Leu
1 5 10 15

Ser Pro Thr Arg Leu Ser Arg Leu Gln Glu Lys Glu Glu Leu Arg Glu
20 25 30

Leu Asn Asp Arg Leu Ala His Tyr Ile Asp Arg Val Arg Ala Leu Glu
35 40 45

Leu Glu Asn Asp Arg Leu Leu Leu Lys Ile Ser Glu Lys Glu Glu Val
50 55 60

Thr Thr Arg Glu Val Ser Gly Ile Lys Ala Leu Tyr Glu Ser Glu Leu
65 70 75 80

Ala Asp Ala Arg Arg Val Leu Asp Glu Thr Ala Arg Glu Arg Ala Arg
85 90 95

Leu Gln Ile Glu Ile Gly Lys Leu Arg Ala Glu Leu Asp Glu Val Asn
100 105 110

Lys Ser Ala Lys Lys Arg Glu Gly Glu Leu Thr Val Ala Gln Gly Arg
115 120 125

Val Lys Asp Leu Glu Ser Leu Phe His Arg Ser Glu Val Glu Leu Ala
130 135 140

Ala Ala Leu Ser Asp Lys Arg Gly Leu Glu Ser Asp Val Ala Glu Leu
145 150 155 160

Arg Ala Gln Leu Ala Lys Ala Glu Asp Gly His Ala Val Ala Lys Lys

[illegible]

329

Lys Arg Leu Glu Val Glu Glu Pro Leu Gly Ser Gly Pro Ser Val Leu
 420 425 430

Gly Thr Gly Thr Gly Gly Ser Gly Gly Phe His Leu Ala Gln Gln Ala
 435 440 445

Ser Ala Ser Gly Ser Val Ser Ile Glu Glu Ile Asp Leu Glu Gly Lys
 450 455 460

Phe Val Gln Leu Lys Asn Asn Ser Asp Lys Asp Gln Ser Leu Gly Asn
 465 470 475 480

Trp Arg Ile Lys Arg Gln Val Leu Glu Gly Glu Glu Ile Ala Tyr Lys
 485 490 495

Phe Thr Pro Lys Tyr Ile Leu Arg Ala Gly Gln Met Val Thr Val Trp
 500 505 510

Ala Ala Gly Ala Gly Val Ala His Ser Pro Pro Ser Thr Leu Val Trp
 515 520 525

Lys Gly Gln Ser Ser Trp Gly Thr Gly Glu Ser Phe Arg Thr Val Leu
 530 535 540

Val Asn Ala Asp Gly Glu Glu Val Ala Met Arg Thr Val Lys Lys Ser
 545 550 555 560

Ser Val Met Arg Glu Asn Glu Asn Gly Glu Glu Glu Glu Glu Glu Ala
 565 570 575

Glu Phe Gly Glu Glu Asp Leu Phe His Gln Gln Gly Asp Pro Arg Thr
 580 585 590

Thr Ser Arg Gly Cys Tyr Val Met
 595 600

<210> 257
 <211> 620
 <212> PRT
 <213> Homo sapien

<400> 257

Met Ser Pro Pro Ser Pro Gly Arg Arg Arg Glu Gln Arg Arg Pro Arg
 1 5 10 15

330

Ala Ala Ala Thr Met Ala Thr Pro Leu Pro Gly Arg Ala Gly Gly Pro
 20 25 30

Ala Thr Pro Leu Ser Pro Thr Arg Leu Ser Arg Leu Gln Glu Lys Glu
 35 40 45

Glu Leu Arg Glu Leu Asn Asp Arg Leu Ala His Tyr Ile Asp Arg Val
 50 55 60

Arg Ala Leu Glu Leu Glu Asn Asp Arg Leu Leu Leu Lys Ile Ser Glu
 65 70 75 80

Lys Glu Glu Val Thr Thr Arg Glu Val Ser Gly Ile Lys Ala Leu Tyr
 85 90 95

Glu Ser Glu Leu Ala Asp Ala Arg Arg Val Leu Asp Glu Thr Ala Arg
 100 105 110

Glu Arg Ala Arg Leu Gln Ile Glu Ile Gly Lys Leu Arg Ala Glu Leu
 115 120 125

Asp Glu Val Asn Lys Ser Ala Lys Lys Arg Glu Gly Glu Leu Thr Val
 130 135 140

Ala Gln Gly Arg Val Lys Asp Leu Glu Ser Leu Phe His Arg Ser Glu
 145 150 155 160

Val Glu Leu Ala Ala Ala Leu Ser Asp Lys Arg Gly Leu Glu Ser Asp
 165 170 175

Val Ala Glu Leu Arg Ala Gln Leu Ala Lys Ala Glu Asp Gly His Ala
 180 185 190

Val Ala Lys Lys Gln Leu Glu Lys Glu Thr Leu Met Arg Val Asp Leu
 195 200 205

Glu Asn Arg Cys Gln Ser Leu Gln Glu Glu Leu Asp Phe Arg Lys Ser
 210 215 220

Val Phe Glu Glu Glu Val Arg Glu Thr Arg Arg Arg His Glu Arg Arg
 225 230 235 240

Leu Val Glu Val Asp Ser Ser Arg Gln Gln Glu Tyr Asp Phe Lys Met
 245 250 255

Ala Gln Ala Leu Glu Glu Leu Arg Ser Gln His Asp Glu Gln Val Arg

331

260	265	270
Leu Tyr Lys Leu Glu Leu Glu Gln Thr Tyr Gln Ala Lys Leu Asp Ser		
275	280	285
Ala Lys Leu Ser Ser Asp Gln Asn Asp Lys Ala Ala Ser Ala Ala Arg		
290	295	300
Glu Glu Leu Lys Glu Ala Arg Met Arg Leu Glu Ser Leu Ser Tyr Gln		
305	310	315
Leu Ser Gly Leu Gln Lys Gln Ala Ser Ala Ala Glu Asp Arg Ile Arg		
325	330	335
Glu Leu Glu Glu Ala Met Ala Gly Glu Arg Asp Lys Phe Arg Lys Met		
340	345	350
Leu Asp Ala Lys Glu Gln Glu Met Thr Glu Met Arg Asp Val Met Gln		
355	360	365
Gln Gln Leu Ala Glu Tyr Gln Glu Leu Leu Asp Val Lys Leu Ala Leu		
370	375	380
Asp Met Glu Ile Asn Ala Tyr Arg Lys Leu Leu Glu Gly Glu Glu Glu		
385	390	395
Arg Leu Lys Leu Ser Pro Ser Pro Ser Ser Arg Val Thr Val Ser Arg		
405	410	415
Ala Thr Ser Ser Ser Ser Gly Ser Leu Ser Ala Thr Gly Arg Leu Gly		
420	425	430
Arg Ser Lys Arg Lys Arg Leu Glu Val Glu Glu Pro Leu Gly Ser Gly		
435	440	445
Pro Ser Val Leu Gly Thr Gly Thr Gly Gly Ser Gly Gly Phe His Leu		
450	455	460
Ala Gln Gln Ala Ser Ala Ser Gly Ser Val Ser Ile Glu Glu Ile Asp		
465	470	475
Leu Glu Gly Lys Phe Val Gln Leu Lys Asn Asn Ser Asp Lys Asp Gln		
485	490	495
Ser Leu Gly Asn Trp Arg Ile Lys Arg Gln Val Leu Glu Gly Glu Glu		
500	505	510

332

Ile Ala Tyr Lys Phe Thr Pro Lys Tyr Ile Leu Arg Ala Gly Gln Met
 515 520 525

Val Thr Val Trp Ala Ala Gly Ala Gly Val Ala His Ser Pro Pro Ser
 530 535 540

Thr Leu Val Trp Lys Gly Gln Ser Ser Trp Gly Thr Gly Glu Ser Phe
 545 550 555 560

Arg Thr Val Leu Val Asn Ala Asp Gly Glu Glu Val Ala Met Arg Thr
 565 570 575

Val Lys Lys Ser Ser Val Met Arg Glu Asn Glu Asn Gly Glu Glu Glu
 580 585 590

Glu Glu Glu Ala Glu Phe Gly Glu Glu Asp Leu Phe His Gln Gln Gly
 595 600 605

Asp Pro Arg Thr Thr Ser Arg Gly Cys Tyr Val Met
 610 615 620

<210> 258
 <211> 237
 <212> PRT
 <213> Homo sapien

<400> 258

Met Ser Pro Pro Ser Pro Gly Arg Arg Arg Glu Gln Arg Arg Pro Arg
 1 5 10 15

Ala Ala Ala Thr Met Ala Thr Pro Leu Pro Gly Arg Ala Gly Gly Pro
 20 25 30

Ala Thr Pro Leu Ser Pro Thr Arg Leu Ser Arg Leu Gln Glu Lys Glu
 35 40 45

Glu Leu Arg Glu Leu Asn Asp Arg Leu Ala His Tyr Ile Asp Arg Val
 50 55 60

Arg Ala Leu Glu Leu Glu Asn Asp Arg Leu Leu Leu Lys Ile Ser Glu
 65 70 75 80

Lys Glu Glu Val Thr Thr Arg Glu Val Ser Gly Ile Lys Ala Leu Tyr
 85 90 95

333

Glu Ser Glu Leu Ala Asp Ala Arg Arg Val Leu Asp Glu Thr Ala Arg
 100 105 110

Glu Arg Ala Arg Leu Gln Ile Glu Ile Gly Lys Leu Arg Ala Glu Leu
 115 120 125

Asp Glu Val Asn Lys Ser Ala Lys Lys Arg Glu Gly Glu Leu Thr Val
 130 135 140

Ala Gln Gly Arg Val Lys Asp Leu Glu Ser Leu Phe His Arg Ser Glu
 145 150 155 160

Val Glu Leu Ala Ala Ala Leu Ser Asp Lys Arg Gly Leu Glu Ser Asp
 165 170 175

Val Ala Glu Leu Arg Ala Gln Leu Ala Lys Ala Glu Asp Gly His Ala
 180 185 190

Val Ala Lys Lys Gln Leu Glu Lys Gly Cys Pro Cys Ser Gln Lys Ser
 195 200 205

Arg Ser His Val Asp Arg Gly Gly Arg Ile Leu Pro Lys His Phe Leu
 210 215 220

Leu Glu Ala Thr Pro Leu Cys Ser Gln Ser Gly Gly Trp
 225 230 235

<210> 259
 <211> 620
 <212> PRT
 <213> Homo sapien

<400> 259

Met Ser Pro Pro Ser Pro Gly Arg Arg Arg Glu Gln Arg Arg Pro Arg
 1 5 10 15

Ala Ala Ala Thr Met Ala Thr Pro Leu Pro Gly Arg Ala Gly Gly Pro
 20 25 30

Ala Thr Pro Leu Ser Pro Thr Arg Leu Ser Arg Leu Gln Glu Lys Glu
 35 40 45

Glu Leu Arg Glu Leu Asn Asp Arg Leu Ala His Tyr Ile Asp Arg Val
 50 55 60

Arg Ala Leu Glu Leu Glu Asn Asp Arg Leu Leu Leu Lys Ile Ser Glu
 65 70 75 80

334

Lys Glu Glu Val Thr Thr Arg Glu Val Ser Gly Ile Lys Ala Leu Tyr
85 90 95

Glu Ser Glu Leu Ala Asp Ala Arg Arg Val Leu Asp Glu Thr Ala Arg
100 105 110

Glu Arg Ala Arg Leu Gln Ile Glu Ile Gly Lys Leu Arg Ala Glu Leu
115 120 125

Asp Glu Val Asn Lys Ser Ala Lys Lys Arg Glu Gly Glu Leu Thr Val
130 135 140

Ala Gln Gly Arg Val Lys Asp Leu Glu Ser Leu Phe His Arg Ser Glu
145 150 155 160

Val Glu Leu Ala Ala Ala Leu Ser Asp Lys Arg Gly Leu Glu Ser Asp
165 170 175

Val Ala Glu Leu Arg Ala Gln Leu Ala Lys Ala Glu Asp Gly His Ala
180 185 190

Val Ala Lys Lys Gln Leu Glu Lys Glu Thr Leu Met Arg Val Asp Leu
195 200 205

Glu Asn Arg Cys Gln Ser Leu Gln Glu Glu Leu Asp Phe Arg Lys Ser
210 215 220

Val Phe Glu Glu Glu Val Arg Glu Thr Arg Arg Arg His Glu Arg Arg
225 230 235 240

Leu Val Glu Val Asp Ser Ser Arg Gln Gln Glu Tyr Asp Phe Lys Met
245 250 255

Ala Gln Ala Leu Glu Glu Leu Arg Ser Gln His Asp Glu Gln Val Arg
260 265 270

Leu Tyr Lys Leu Glu Leu Glu Gln Thr Tyr Gln Ala Lys Leu Asp Ser
275 280 285

Ala Lys Leu Ser Ser Asp Gln Asn Asp Lys Ala Ala Ser Ala Ala Arg
290 295 300

Glu Glu Leu Lys Glu Ala Arg Met Arg Leu Glu Ser Leu Ser Tyr Gln
305 310 315 320

Leu Ser Gly Leu Gln Lys Gln Ala Ser Ala Ala Glu Asp Arg Ile Arg
325 330 335

Glu Leu Glu Glu Ala Met Ala Gly Glu Arg Asp Lys Phe Arg Lys Met
340 345 350

Leu Asp Ala Lys Glu Gln Glu Met Thr Glu Met Arg Asp Val Met Gln
355 360 365

Gln Gln Leu Ala Glu Tyr Gln Glu Leu Leu Asp Val Lys Leu Ala Leu
370 375 380

Asp Met Glu Ile Asn Ala Tyr Arg Lys Leu Leu Glu Gly Glu Glu Glu
385 390 395 400

Arg Leu Lys Leu Ser Pro Ser Pro Ser Ser Arg Val Thr Val Ser Arg
405 410 415

Ala Thr Ser Ser Ser Ser Gly Ser Leu Ser Ala Thr Gly Arg Leu Gly
420 425 430

Arg Ser Lys Arg Lys Arg Leu Glu Val Glu Glu Pro Leu Gly Ser Gly
435 440 445

Pro Ser Val Leu Gly Thr Gly Thr Gly Gly Ser Gly Gly Phe His Leu
450 455 460

Ala Gln Gln Ala Ser Ala Ser Gly Ser Val Ser Ile Glu Glu Ile Asp
465 470 475 480

Leu Glu Gly Lys Phe Val Gln Leu Lys Asn Asn Ser Asp Lys Asp Gln
485 490 495

Ser Leu Gly Asn Trp Arg Ile Lys Arg Gln Val Leu Glu Gly Glu Glu
500 505 510

Ile Ala Tyr Lys Phe Thr Pro Lys Tyr Ile Leu Arg Ala Gly Gln Met
515 520 525

Val Thr Val Trp Ala Ala Gly Ala Gly Val Ala His Ser Pro Pro Ser
530 535 540

Thr Leu Val Trp Lys Gly Gln Ser Ser Trp Gly Thr Gly Glu Ser Phe
545 550 555 560

336

Arg Thr Val Leu Val Asn Ala Asp Gly Glu Glu Val Ala Met Arg Thr
565 570 575

Val Lys Lys Ser Ser Val Met Arg Glu Asn Glu Asn Gly Glu Glu Glu
580 585 590

Glu Glu Glu Ala Glu Phe Gly Glu Glu Asp Leu Phe His Gln Gln Gly
595 600 605

Asp Pro Arg Thr Thr Ser Arg Gly Cys Tyr Val Met
610 615 620